

European defence in the melting-pot

Decisions due in the next few months could help to stabilize the defence of Western Europe at a less costly level as well as help Mr Gorbachev survive troubles such as those in the Baltic states.

REASONS why nobody can quite tell the future of European defence hardly need enumeration, but there is an interesting array of them:

■ A new president, Mr George Bush, will be installed in the White House on 20 January, and while there is nothing to suggest a break with US policies of the past eight years, it may be significant that Mr Bush himself has been non-committal on the subject. The new administration, in seeking both to cut the federal deficit and to satisfy long-standing conservative complaints, could well insist that Western Europe should bear a larger fraction of the cost of, perhaps even of the responsibility for, its own defence.

■ Second, Western Europe is still bemused by the arrival of Mr Mikhail Gorbachev in power in the Soviet Union, is impressed that there has already been one significant arms control agreement (on weapons of intermediate range), that there is another (on strategic weapons) waiting for Mr Bush's decision about the future of the Strategic Defense Initiative (SDI) and that serious negotiations on yet a third (on conventional forces) are due to begin in the spring. That is why it is taking the West Germans an age to decide whether to accept, early in the 1990s, improved substitutes for the US Lance missiles now sited in West Germany, whose supposed military role is the interdiction of massive conventional forces attacking from the East.

■ Eastern Europe, once regarded as indistinguishable from the Soviet Union, no longer seems like that. Taking liberal-sounding Gorbachev perhaps too literally at his word, the Baltic states have been making declarations of what might be called separate-ness, while Hungary, traditionally in the van of eastern change, seems bent on much more of the same. Even Poland and Czechoslovakia now seem more flexible than a few years or even months ago, while politicians like President Mitterrand of France, no doubt anxious to gain what they can from changing circumstances, have taken to speaking about the restoration of traditional links with the Balkan states.

Back burner?

So does all that not imply that European defence, meaning the defence of Western Europe, is well on the way to becoming a non-issue, something for the back burner? Sadly, no, because:

■ Signs that the Russian Republic may now be less in charge of events on its periphery than a few years ago are not quite the general comfort in the West that they would have been in, say, the 1970s. Instead, they raise in the minds of Western military planners the uneasy prospect that centrifugal tendencies in the Baltic states may precipitate the unthinkable — Mr Mikhail Gorbachev's own collapse, followed by that of *perestroika* (always a high-risk strategy) and by the return of implacability in the East.

Not for the first time, not all these tendencies are harmoniously in accord. The last and, most often, the unspoken consideration is dominant, but can be argued two ways. One, the most common, is that the possibility that the new liberalism may not be permanent predicates a continuation into the 1990s of the defence arrangements of the 1970s, when there was not even an agreement on nuclear weapons of intermediate range. Specifi-

cally, this would imply continued reliance on nuclear interdiction of conventional attacks from the East — a view that could only be strengthened by US insistence that Western Europe should learn to stand on its own feet (for nuclear weapons cost only money; strengthening conventional forces entails both monetary and political costs). More generally, this interpretation leads to the conclusion that Western Europe must be armed (to defend itself) to the teeth until Mr Gorbachev can demonstrate that his concept of how the Soviet Union should be run has come to stay.

Opposite interpretation

That is why more attention should be given to the opposite interpretation of events in the Baltic states and elsewhere in Eastern Europe: that this is eminently the time for striking deals between East and West. The negotiations on conventional forces arranged for next year will evidently be crucial both to East-West relations and the Western European response to continuing pressure from the United States. The snag is that they are bound to take an age and a day. Much, in any case, will hang on the review of policy by the North Atlantic Treaty Organization, also due in the spring. The best outcome would be a declaration that Western Europe and its allies will trade reductions of interdiction weapons (from nuclear artillery to Lance replacements) for progress in the reduction of conventional forces. In retrospect unwisely, the West insisted that these weapons should be exempted from the impending negotiations on conventional weapons.

If that is indeed what happens in the West in the next few months, several stabilizing influences will follow. First, Mr Gorbachev will have something more to show for his risky willingness to negotiate. Second, it will then seem to his own colleagues a less serious matter if there should persist centrifugal tendencies in peripheral states. (Conversely, progress towards the reduction of interdiction weapons would assist centrifugal forces.) Third, the West German government will be partially relieved of its well-publicized agony about the prospect that nuclear weapons used for interdiction will as often as not fall on West German soil. Fourth, and no less important, these same steps would resolve the potential tension between the United States and Western Europe over their mutual obligations. In short, there is a way in which the apparent circle of contradictions might be squared. □

Student debtors rally

British students have been protesting at the proposed ending of their grants. They have cause.

CAPITAL cities are incomplete without occasional conflicts between students and the authorities, so that the British government must be commended for having provided the occasion for last Thursday's near-riot in central London. Many years have passed since students and mounted police were previously in conflict during an evening rush-hour. (It is strictly coincidental that Hungarian students were also up in arms last week.



eventually burning a copy of the education act in the main hall of the Oeþtvops Lornad University in Budapest: they may be thought to have even more substantial grievances.) The occasion, in London last week, was the government's plan to replace by loans part of the maintenance grant to which students in higher education are at present entitled.

Most previous essays in this direction have foundered on the calculation that the short-run cost of loans is almost certain to be much greater than that of maintenance grants. The scheme devised by the Department of Education and Science avoids some of these fiscal snags. Loans will only gradually replace public subventions, with the result that there will be extra public costs of between £100 million and £200 million in the remainder of this century, followed by increasing public savings. But the scheme ends the principle of entitlement, which predictably brought the students onto the streets. The scheme is also intended to prevent students from supplementing their incomes by claiming benefits from the social security system. After a transition period whose length will be inversely correlated with the inflation rate, British students will, on this plan, be more than half dependent on loans.

Fiscally, the scheme is tidy and politically expedient. The real value of grants to individual students has been declining steadily for the past quarter of a century, chiefly because successive governments have not fully compensated students for inflation. The slack has been taken up by contributions from students' parents, chiefly by the devices of neglecting fully to index against inflation the incomes at which parents are required to pay certain percentages of the estimated total cost and partly because students' parents have progressively become more prosperous. The result is that, by the academic year 1986-87, when the average nominal maintenance grant for students was £2,018, students' parents were on the average expected to pay roughly £800 while the average student's income was supplemented by roughly a third of the value of the nominal grant, to a total of £2,686, of which roughly £250 (on the average) came from the social security system.

What the British government now says is that the time has come to end student dependence on the social security system and even on taxpayers, "many of whom do not share the advantages that students have once they graduate", and that parental contributions should similarly be decreased. The plan is that, from September 1990, both the government's and parents' contributions to student maintenance should be frozen in cash terms (and thus made hostages to inflation) and that students should be entitled to an interest-free loan (an average of £420 from next September) which will be increased steadily with inflation until it exceeds the sum of the nominal contributions of the government and parents. On recent form (British inflation this year, standing at more than 6 per cent, exceeds the forecasts put out in March), that may happen before the end of the century.

Strictly fiscal

Collectively, the defects of these proposals is that they are strictly fiscal. The British government seems simply to have applied to the system of higher education the housekeeping principles for which it is famous, and in particular its belief that all kinds of public expenditure are bad, without having paid attention to the underlying educational issues. It is astonishing that the scheme should have been made public, in a document described as a "white paper" and thus supposedly a policy statement, without saying much about repayment terms and before the government had established whether the commercial banks, to which it looks for the administration of the scheme, would be prepared to help. (If the British government carries on like this, it will help to reinforce the now-substantial opinion that its electoral strength has gone to its head, making it high-handed.) In reality, the cost to the government (or to British taxpayers) has increased to £830 million a year, or roughly a

fifth of public spending on higher education, and has at various times in the past decade been given as a reason for freezing the number of student places in the university sector. Nobody has bothered to ask whether 20 per cent is too much, too little or just right.

The issue most seriously neglected is that of access to higher education. The British government is rightly proud of its improvement of the economic framework during the past decade (last week's deterioration of the external trade balance notwithstanding) but has perversely never considered the contribution of higher education to the national well-being except in the most negative way, by means of a continuing search for means of reducing its cost to public funds. The new loans scheme is squarely in this tradition. (It has the further political advantage of placating middle-class families resentful that the burden of student support increasingly falls on them.) But what the British government should be asking is whether, even under present arrangements, the proportions of young people acquiring further skills after leaving secondary education are sufficient for Britain's economic needs and whether the loans scheme will help or hinder.

Gloomy outlook

The outlook is gloomy. Although the proportion of young people remaining after secondary school in some form of higher education is now 31 per cent, only a third of these people follow degree courses at universities and polytechnics. Yet much of British industry is hamstrung by long-standing shortages of skill — for example in computer science — and newly emerging shortages, in engineering of most kinds; nobody knows the extent to which its creativity is hampered because skilled people are not in surplus. Yet the assertion in the white paper that experience elsewhere shows that students' decisions about higher education "do not depend on financial considerations alone" (which, platitudinously, may well be true) takes no account of the likelihood that many intending students' perceptions that higher education will saddle them with insupportable debts will impel them towards other career paths. The government's calculation that the financial return to society of public spending on higher education (between 5.0 and 8.0 per cent a year) is less than the rate of return to students on their investment (income forgone and the value of maintenance grants) of between 22.0 and 27.5 per cent a year, and that the disparity should be reduced, is strictly meaningless. By that test, virtually all public expenditure would be declared invalid.

Although last week's protest may have been sustained by generalized discontent, it can only mistakenly be dismissed as such. There is every likelihood that the loans scheme will further diminish the recruitment to degree courses of young people, especially those from less prosperous families. Is that what a revitalized British economy really needs? And there is little prospect that the three funds, each of £5 million, which the government has offered as a means of meeting students' needs, will substantially offset that difficulty.

Even at this late stage, the government would be well-advised to start from the other end of a problem in whose solution everybody has an interest. First, on its own principle that social benefits should be targeted towards those who need them, it should introduce a scheme for awarding generous bursaries to degree students whose parental income is below the threshold at which parental contributions are called for; properly designed, such a scheme could establish the elasticity of demand for student places. That should be coupled with an undertaking that universities and polytechnics should be enabled to expand to meet reasonable demands. Then, but only then, could the government introduce a version of its loans scheme without running the risk of keeping out of higher education the young people on whose skills its own successors' well-being will depend. To ask for this is not to make a theoretical argument in favour of wider access, but is merely to urge prudence. □

UK government agrees to anonymous HIV testing

- Scientific pressure forces decision
- Legal and ethical problems denied

London

INTENSE lobbying has persuaded the UK government to agree to the involuntary and anonymous screening of blood samples for human immunodeficiency virus (HIV) infection. The need for information on the prevalence and spread of HIV has overridden objections on both ethical and practical grounds to screening. The Medical Research Council (MRC) has been commissioned by the Department of Health to put forward a plan of action.

Announcing the decision last week, the Secretary of State for Health, Mr Kenneth Clarke, confirmed that the government also supported screening on a voluntary and named basis. Two such studies will start in Scottish antenatal clinics this week, and the MRC is being asked to design others. "Many important matters of detail as to scale and groups to be tested need to be determined", said Clarke.

Large-scale voluntary testing was first approved by the UK government in June, following the recommendations of a working party under Dr Joe Smith, director of the Public Health Laboratory Service (see *Nature* 333, 486; 9 June 1988). But the working party came to the view that involuntary testing was not needed because the data required could be obtained by voluntary testing and "there are legal, ethical and scientific methodological problems" in undertaking such studies. Last week, Kenneth Clarke declared that the government considers there is neither legal obstacle nor ethical objection to testing for the presence of HIV antibodies in blood samples taken for other purposes, as long as the sample's donor is not identifiable.

Remaining problems of methodology are to be solved by the MRC, which has been asked to come forward with proposals within three months, and has been given £1.7 million for research into the spread of HIV infection, for the financial year 1989-90. The MRC Committee on Epidemiological Studies of AIDS has already been considering, in principle, the best design for involuntary and anonymous screening, and will now have to flesh out the details. Committee chairman Dr Nick Day says that there should be no difficulty in producing detailed and costed proposals within the allotted time, and that some studies could begin before then, subject to peer review.

Day's predecessor, Sir Richard Doll, was among the signatories of a letter in

The Lancet almost exactly a year ago that argued strongly in favour of both voluntary and involuntary anonymous screening, and criticized the House of Commons Social Services Committee on AIDS for having rejected such screening in the face of scientific and medical opinion. The letter said the committee's decision reflected the arguments of only a handful of witnesses, and was based on statistical misunderstanding, unsupported ethical assertions and a failure to appreciate that anonymous testing can still provide prevalence data for specific risk groups.

A major task for the MRC will be to design protocols that allow such risk factors to be gathered. Clearly that will be easy in studies of blood samples from clinics that specialize in sexually transmitted diseases, as the relevant questions will have been asked as a matter of routine. At the other end of the spectrum, blood samples taken in, for example, general surgical wards of hospitals are unlikely to be routinely accompanied by information on risk factors. As a minimum, all samples will be accompanied by information on the age and sex of the patient, and the district and type of medical facility from where the sample was drawn.

Precisely what has caused the government belatedly to agree to anonymous screening is unclear. One factor, however, must be a report prepared by statistician Sir David Cox, warden of Nuffield College, Oxford. The report, which is due to be published shortly, explicitly recommends anonymous testing. Naturally, Cox welcomes the government decision and is pleased by the fact that the MRC is to design the testing programme. There is a serious job to be done, he says, particularly if the maximum information about the individuals from whom the samples have come is to be obtained without allowing identification of individuals and while avoiding any element of opting out.

The problems of opting out encountered in voluntary schemes of HIV testing are likely to have been another factor that influenced the government's decision. The proportion of people giving consent is variable but has been falling, says Day.

Overall, the decision has been taken because of the need to obtain more accurate information on the number of people now infected with HIV in Britain — 50,000 is the 'guesstimate' — and the spread of infection, so that healthcare needs can be predicted and planned. **Peter Newmark**

Computer files on AIDS carriers

Paris

AFTER almost a year of debate, the French national council for the freedom of information (CNIL) has approved an epidemiological study of AIDS using a computer database. The study, to be carried out by researchers from the national institute of health and medical research (INSERM) will collect data from the 23 national AIDS information and care centres. Information on the mode of transmission of the virus, together with encrypted personal details, will be recorded from seropositive individuals who have given their written permission.

Peter Coles

Canadian excellence

Washington

WHILE a fiercely fought election campaign in Canada has just ended in victory for free trade and the Progressive Conservative party of Prime Minister Brian Mulroney, another, quieter battle is shaping up over who will be selected for one of the ten 'Networks of Centres of Excellence' that lie at the heart of the government's scheme to add new life to Canadian science.

The response to the Centres programme has been staggering. Since the May decision (see *Nature* 333, 722; 23 June 1988) to spend C\$240 million over five years on research networks linking universities, industry and government, 238 letters of intent to submit proposals for networks have been received. That response gives some measure of how many Canadian researchers are hungry for research funds; despite Canada's wealth, spending for research measured as a percentage of gross national product places Canada among the developing countries.

Some applicants will now look at the competition and decide to drop out before the final deadline for full proposals expires on 30 November. Others will see possible marriages of convenience and combine with rival groups to build bigger, stronger proposals. But the government will have to reject 90 per cent of the proposals.

This week, the final proposals will go to Canada's three research councils which will manage an international peer review process to assess them. The assessments will then be passed to an independent National Advisory Board on Centres of Excellence, headed by Allelix president John Evans and University of Montreal rector Gilles Cloutier, for the preparation of final rankings. With that done, the government will take over and make choices early in the spring. Plans are for the ranked list to be published, in order to lessen any temptation for the government to meddle. There is still widespread concern that the choices could be made to satisfy regional, rather than scientific, interests.

Alun Anderson

NIH cuts bring in hard times for biomedical researchers

Berkeley

BIOMEDICAL training in the United States is in for hard times following a decision by the National Institutes of Health (NIH) to reduce the number of predoctoral and postdoctoral positions it supports. In some institutes it may become impossible to finance new training grants or renew those due to expire.

Training grants provide stipends and tuition expenses for predoctoral graduate students and stipends for postdoctoral trainees, as well as 'supply funds' that many departments use to finance seminar series and courses. The squeeze will also reduce the number of individual postdoctoral fellowships.

Close to 1,000 posts were put at risk after Congress refused to appropriate money to pay for an increase in fellowship stipends decided upon by NIH. But some say the crisis has been brewing for years, as funding for NIH training programmes has failed to keep up with the rise in the cost of living and tuition expenses.

The stipend increase "exacerbated a very bad situation", says William Pitlick, research training officer at NIH. But NIH could not ignore criticism by university administrators that the stipend level had not been raised for four years.

According to Pitlick, postdoctoral training grants were intended to encourage people trained as doctors to obtain research experience at the postdoctoral level. But the grants were thousands of dollars less than salaries paid to medical residents and created little incentive for good doctors to enter research. Graduate stipends of \$6,500 were also unreasonably low, and had to be supplemented substantially by universities. As a result of the increase, graduate stipends will rise from \$6,500 to \$8,500, and a third-year postdoctoral fellow who previously would have earned \$21,000 will now earn \$25,000 per year.

The National Institute of General Medical Sciences (NIGMS), which finances two-thirds of NIH-sponsored predoctoral training grants, must cut 400 training positions, 300 of them from predoctoral programmes. The Medical Scientist Training Program (MSTP) which funds MD/PhD students will lose 155 of its present 725 positions. Institute director Ruth Kirschstein says it is uncertain whether there will be money to fund any new training grants or renew those at the end of their five-year grant period.

Murray Goldstein, director of the National Institute for Neurological Disorders and Stroke (NINDS), sent a ripple of panic through the neuroscience community last month when he informed

training grant holders there would be no funds to renew five-year grants. The institute would normally fund 15 to 18 grants, Goldstein says, but this year, "twenty-one training grants are approved for our January meeting, to compete for no dollars".

Among those ill-fated renewal applications are training grants from the neuroscience graduate programmes at the Harvard, Stanford, Columbia and Yale medical schools. Story Landis, of the neuroscience programme at Case Western Reserve University, which also stands to lose its training grant, has started a grassroots movement among neuroscientists to lobby Congress for an increase in funds. The loss of grants will "eviscerate the training environment", says Landis. Not only will it force departments to take fewer students and post-docs, she says, but it will eliminate funds used to support seminar series and graduate student recruitment.

Goldstein agrees that training grants serve an important role, but NINDS training grants support mostly postdoctoral trainees, and Goldstein says that the

money is better spent on individual fellowships, where it can buy more positions. "Given the very difficult situation, do we give that institution money to enrich their programme, at the cost of a fellowship?" At NINDS, for the time being, the answer is no. Training grants that are not funded this year may be re-submitted next year, when there will be some funds available. But the competition will be stiff.

The crisis is not temporary according to Pitlick. Although Congress has always funded stipend increases in the past, the current preoccupation with the national budget deficit has made that unlikely. "At the time we implemented the stipend increase, we discussed the distinct possibility that funds wouldn't be available for it", Pitlick says. "We knew it would cost us trainees if we didn't get more money, and that's what's happening." With funds as short as they are, many of those affected would prefer that NIH forgo the stipend increase rather than cut positions. Says NIGMS's Kirschstein, "In this day and age of concern about whether the United States is competitive with other countries in biomedical and particularly biotechnological research, we need to train more people, and we are training fewer."

Marcia Barinaga

More disruption for UK research councils

London

THE research councils for the natural environment (NERC) and for agriculture and food (AFRC) will go into a new year battered and bruised after several years of reductions in public resources. There is little sign that the squeeze is coming to an end and more disruption is on the horizon, with a possible merger between the two councils to form a Natural Resources Research Council (NRRC).

The House of Lords select committee on science and technology has recommended such a merger in an interim report on the future of agricultural and food research; the recommendation is welcomed by both councils, though more so by the NERC which recommended the merger. The AFRC, instead, strongly advocated a single council for non-medical biology, which would be called the Biological Resources Council (BRC), bringing together the biological activities of the NERC, the AFRC and the Science and Engineering Research Council (SERC). But the SERC itself is opposed to this, arguing that at present it can encourage collaboration between life scientists and physical and chemical scientists and mathematicians. And many believe it would be a step backwards if studentships in biological sciences were to be handled separately from those in physics, chemistry and engineering.

The Lords committee decided against the formation of a BRC, which, it said,

could only have advantages if the council were also responsible for the biological research carried out by the Medical Research Council (MRC); but because of the special character and needs of medical research, the committee decided that the MRC should continue its separate existence. Meanwhile, the formation of a NRRC would remove some of the inappropriate divisions of responsibility between the two councils. The resulting organization will be stronger and more effective than its constituent parts, says the committee. It would be comparable in size to the SERC and the MRC, and so could lead to a better balance within the research council structure. If the government accepts the Lords' advice, following the committee's full report next spring, the councils will then begin to discuss exactly how it should be carried out. The NERC urges that it should be a gradual development; the council's new chairman, Professor John Knill, is cautious about the move and keen that the NERC does not lose its distinct role as the council responsible for the environment, that role being one the government now supports and one with which the public can identify.

If the merger is to go ahead, then both councils will insist that funds for restructuring do not come out of research funds. In their annual reports, both refer with dismay to the reduction in public funding which has led to less research and to redundancies.

Christine McGourty

Soviet Union puts a Frenchman and two Russians into space

Paris

ON Saturday evening (26 November), a Soviet Soyuz-TM launcher sent a Frenchman, Jean-Loup Chrétien, and two Soviet cosmonauts, Alexandre Volkov and Serguei Krikalev, on a two-day voyage to the Mir space station, where they will stay for three weeks. For Soviet cosmonauts, this kind of mission has become almost routine — three have spent more than 320 days aboard Mir — but it will provide European space scientists with their first chance to collect physiological data over a prolonged period in preparation for the European Space Agency's own projected manned flights aboard the Hermes shuttle at the end of the century.

Chrétien, aged 50, is Europe's most travelled astronaut. In 1982, he took part in another Soviet mission to the Salyut-7 space station. On the present mission, Aragatz, he will join one of his Soviet

colleagues in a 'space walk' to assemble an antenna and experimental instruments designed to collect data on the different kinds of stress to which external structures are subjected in space. But the principal interest for European scientists lies in two biomedical studies on man's physiological response to prolonged weightlessness, code-named PHYSALIE and VIMINAL.

These two studies have been designed by Alain Berthoz, a neurophysiologist at the Centre National de la Recherche Scientifique (CNRS) laboratory for neurosensory physiology, in collaboration with the Moscow Academy of Science, to investigate perceptual-motor adaptation to microgravity conditions. PHYSALIE will collect data on postural adaptation and changes in muscle response; VIMINAL will look at visual-manual manipulation of objects and orientation.

On the eve of the flight, Hubert Curien, the French research minister, who was president of the national centre for space research (CNES) in 1980, lent his support to France's controversial backing of manned space missions, in an article in the national daily newspaper, *Le Monde*. The value of manned missions was contested in a report by the French Academy of Sciences (see *Nature* 333, 2: 12 May 1988), on the grounds that astronauts can perform no useful functions that could not better be carried out by robots. But for Curien, not to participate now in manned missions would make Europe a "second-rate continent" dependent upon the "regal gift of know-how gained by our two major partners [the Soviet Union and the United States] without us".

Peter Coles

New ESA project

Paris

THE EUROPEAN Space Agency (ESA) has chosen the Cassini/Titan probe mission as its next space science project, six years after it was first proposed. Four other missions were in competition for a vacant slot in ESA's Horizon 2000 long-term space science plan (see *Nature* 336, 98: 10 November 1988). The final decision was made at the science programme committee's meeting on 24 November in Paris.

Cassini is a joint mission, in collaboration with the US National Aeronautics and Space Administration (NASA), to explore the atmosphere of Saturn's moon, Titan. Titan is the largest moon in the Solar System, and its nitrogen-rich atmosphere could contain a variety of pre-biotic molecules, providing a model for the origins of life on Earth.

NASA will provide the Saturn orbiter, which will circle the planet for four years, passing over Titan more than 30 times and coming to within 1,000 km of its surface. ESA's Titan probe, named Huyghens, will be released into the moon's atmosphere as soon as the orbiter reaches Saturn, slowing to an impact velocity of 5 metres per second, while the orbiter relays data from the probe to Earth.

Although ESA scientists have contributed individually to other planetary exploration missions, the Giotto mission to comet Halley is the only ESA project to have explored a body in the Solar System other than Earth. The choice of Cassini, due for launch by NASA in April 1996, marks ESA's intention to become directly involved with planetary exploration.

Peter Coles

UK public says 'no, thanks' to nuclear waste

London

As the search continues in Britain for a deposit for the nuclear industry's low- and intermediate-level nuclear waste, the public have overwhelmingly rejected one of the options being considered: a repository under the sea-bed which is accessible from a structure like an oil-rig, based off-shore. The option gained little support because of concern over radioactive contamination of the sea. The attitude of the public to the plans of United Kingdom Nirex Limited, the industry's waste disposal body, was assessed by the University of East Anglia's environmental assessment unit, in response to a document circulated by Nirex last year for public discussion. Though the sub-sea-bed option was not popular with the public, Nirex does not rule it out.

Nirex proposed three deep disposal options: an underground depository inland; a depository under the sea-bed, with access from land, and a depository under the sea-bed, with access from an off-shore structure. In response, a large number of people did not specify a choice, rejecting outright any site in their area. And there was strong support for on-site storage of waste, except from people in areas with existing nuclear installations; there were some responses in favour of deep disposal from the public near the nuclear stations at Dounreay in Caithness, Scotland, and at Sellafield in Cumbria. It is likely that if Nirex does select a site — and it aims to do so by February next year — then it will be one of these.

Christine McGourty

British government promotes atmosphere of concern

London

IN an attempt to prove to the electorate that its conversion to environmentalism is real, the British government is to hold a world conference in March next year on the protection of ozone in the stratosphere. Announcing the conference last week, Mr Nicholas Ridley (right), Secretary of State for the

Environment, said that countries such as India and China had to be persuaded to sign the Montreal Protocol. Their lack of action to protect the ozone layer was due to a lack of information about ozone-benign products and technologies, he said; the aim of the conference will be to provide such information. In April 1990, Britain will host the second meeting to review the Montreal Protocol; the first will be held in Finland in April next year. Lord Caithness (far right), the environment minister, is currently trying to persuade other European Community ministers to ratify the protocol.

At the same time, the race to produce replacements for the chemicals that destroy stratospheric ozone stepped up with the announcement by Britain's biggest chemical company, Imperial Chemical Industries, of the construction of two £30 million commercial plants to produce HFC-134a — a hydrocarbon not containing chlorine. They will be ready for operation in 1991.

Christine McGourty



New vaccine and initiative mean end of rabies in sight for Europe?

London

EUROPE's largest field trial so far of a genetically modified organism has been launched. At the end of October, a 435-square-kilometre tract of land in southern Belgium was baited with a recombinant DNA vaccine against fox rabies. Within two years, it should be clear whether the vaccine is better able to immunize foxes against rabies than the traditional attenuated viral vaccine, but before then the European Commission hopes to have launched a Community-wide drive to vaccinate all foxes in areas where there is rabies and thus to eradicate the disease from Europe.

The Belgian trial will test a vaccinia virus modified to contain a rabies virus gene produced by the French companies Transgène and Rhône-Mérieux. Unlike



Symbol of the European Commission's community-wide drive to control rabies in foxes.

attenuated viruses, the recombinant vaccinia virus can immunize all relevant mammalian species and is not toxic to any of them, according to Professor Paul Pierre Pastoret of the University of Liège, who is in charge of the trial. But because the vaccine is a genetically modified organism, testing it in the field has required extensive approval, ultimately from the National Council of Health, an advisory body to the Belgian Ministry of Health.

The test area has been chosen because it has the lowest density of inhabitants in the country combined with a high incidence of rabies: in the past five years, 148 rabid animals, mainly foxes and cows, have been recorded. Forestry rangers have now distributed 15 baits containing the recombinant DNA vaccine in every square kilometre of the test area apart from a small area that will be studied more intensively by Pastoret and his colleagues.

The success of the test will be measured chiefly by the extent to which foxes in the area become immunized against rabies virus as a result of eating the bait. This will require shooting a number of foxes and trapping others. Additional tests will monitor the uptake of bait and any signs of vaccinia virus lesions. Other domestic and wild mammals will also be monitored.

While the new vaccine is under test in

one locality in southern Belgium, the remainder of the area has recently been baited with attenuated rabies virus. The plan is to eradicate rabies from the whole of the country, as the Swiss have done.

More ambitiously, the European Commission has proposed a plan of action that would remove the disease from the whole Community. If approval is forth-

coming, Community-wide vaccination could be under way next year. To prevent the spread of rabies back into the Community from the east, it would be necessary either to persuade countries on the eastern border also to vaccinate or to maintain a permanently vaccinated zone 60 kilometres wide along the border.

A particular spur for action is that all barriers to trade within the Community are supposed to be dismantled by 1992; one such barrier is the six-month quarantine presently imposed by the United Kingdom and Ireland.

Peter Newmark

West Germany voices objections to European genome project

Munich

THE European Community (EC) is preparing a positive response to the international enthusiasm for mapping and sequencing the human genome, but is having to contend with objections of an ethical character from West Germany.

The community may begin with a programme called "Predictive Medicine" costing a modest 15 million ECU (1 ECU = \$ 1.20) as early as next spring. The first objective is to construct a high-resolution gene map. Independently, the Paris institute CEPH (Centre d'Etude du Polymorphisme Humain) has already begun identifying polymorphisms in the DNA of 40 families. The EC would like the number of families to be increased to 60 and would also like to set up a European network for the cloning and analysis of DNA, together with a parallel computing network to analyse the data.

The second goal is to construct libraries of DNA clones and to store and distribute clones free of charge to European laboratories interested in matching genetic material to the cloned fragments. Finally, the EC would like to improve "advanced genetic technologies" and spread both the technology and the practical knowledge more evenly through the community.

West German parliamentarians do not object to the substance of the programmes, but rather to their intentions, which they claim are based on eugenic principles similar to those of the Nazi movement. Angry words exchanged between Bonn and Brussels bode ill for future agreement. Community officials have been accused of advocating a new form of "racial hygiene" while West Germans have been scolded for "whipping up a misunderstanding". The programme itself has been criticized from other quarters as meagre compared to those being discussed in Japan and already underway in the United States. But an EC official said that the amount committed "does not mean we think that 15 million ECU is all this is worth".

The medical justification given in the proposal has generated most of the controversy. Genetic diseases are said to be "distressing" and "socially very expensive", so that the possibility of relieving or preventing them is promising. The proposal has also given offence by declaring that the programme will strengthen European industry in a "potential European market for DNA probes ... of 1,000 - 2,000 ECU a year".

But, the proposal goes on, the use of genetic information "does raise ethical questions". A study group has been created with a West German, Ernst-Ludwig Winnaker, as its chairman to investigate the broader implications.

In a joint statement on 8 November approved by all the major parties, the Research and Technology Committee of the Bundestag demanded changes in the EC proposal. The committee's strongest objection concerned the way knowledge gained from screening for inheritable diseases could be implemented. Without modifying the human germ line, the critics contend, transmission can only be "prevented" by persuading or even forcing people not to reproduce; a version of "eugenic justification" the committee felt was inappropriate.

This issue will be debated in the plenum of the Bundestag later this month. But the most the parliament could do would be to call upon the West German government to oppose the programme in Brussels.

Ironically, West German research organizations are proceeding apace in setting up a coordinating body for West German research into the human genome. The research council Deutsche Forschungsgemeinschaft (DFG) already has a Priority Programme in "human genome analysis with molecular biological methods", though no West German group has attained international recognition in genome analysis and mapping. The DFG and the Research and Technology Ministry are trying to decide what to do next and how much to invest.

Steven Dickman

Australia reconsiders claims to Antarctic mineral resources

Sydney

The recently concluded draft Convention on the Regulation of Antarctic Resources is forcing Australia to rethink its claims to large areas of the Antarctic continent. Senator Gareth Evans, the Minister for Foreign Affairs and Trade, invited public debate on the convention when he tabled it in parliament on 23 November.

The draft convention was established at a meeting of 16 nations, including Australia, in Wellington on 2 June after six years of negotiations (see *Nature* 333, 588; 16 June 1988).

If ratified, the convention will create a Mineral Resources Commission which would review the environmental impact of any proposed activity and decide whether it would be allowed to proceed. If an area is opened for mining, a regulatory committee would issue licences and monitor compliance with conditions imposed.

According to Senator Evans, the convention would establish a legal framework to protect the environment should mineral activity go ahead. But the convention was formulated on the understanding that Antarctica would stay closed to mineral exploitation "unless a specific decision was taken to open it". He stresses that "contrary to some misconceptions, the aim of the negotiation was not to encourage minerals activity".

Environmental groups would prefer that Antarctica remain untouched. Greenpeace would ban all mining activity and declare the continent a world park. In the United Nations, developing nations have condemned the convention, saying that it would lock up the mineral wealth of the Antarctic for a select group of privileged nations. They argue that Antarctica is the common heritage of humanity, and that all nations should participate in its management and share in its natural riches.

In a letter leaked to the press, Australia's Treasurer Paul Keating urged his cabinet colleagues to be cautious in ratifying the convention because it is a threat to Australian sovereignty over the Antarctic territory it has long claimed. As one of the signatories to the 1959 Antarctic Treaty, Australia would automatically become a member of the commission, a position it would have to share with 19 other nations. Any regulatory committee managing territory presently claimed by Australia would comprise ten nations. But under the terms of the convention no territorial claimant would receive any royalties from mining operations.

According to Keating, that concedes Australia's economic claim to 42 per cent of the continent "for virtually nothing"

and undermines Australia's claim to sovereignty in the Antarctic.

As Australia's territorial claim is so large, the convention cannot come into effect unless it is ratified by Australia. In the near future Australia is bound to come under political pressure from other partners to the convention who are anxious to press ahead and see it ratified.

Charles Morgan

Historical hut in snow danger

Sydney

AUSTRALIAN scientists fear that an important historical monument in Antarctica may soon be lost. Lawson's Hut, the oldest scientific base in the Antarctic, was built in 1912 and used by the explorer Sir Douglas Lawson. Wind-driven snow has blasted away the walls of the hut and reduced them to a paper-thin shell.

In 1962, the first expedition to visit the hut since Lawson's party left in 1914 found that the Baltic pine used to construct the building had been worn away to half its original thickness. In 1985 the wood was down to one quarter of its original thickness. Three attempts to reach the hut since then have failed because of harsh weather.



Dr Bill Burch, a member of the 1962 expedition, fears that an attempt to restore the hut in 1985 might have done more harm than good. Ice and snow packed inside the building was dug out, leaving the building in danger of falling apart.

The hut is listed as a protected building in the Antarctic Treaty and included in Australia's National Estate Register. Burch feels that the hut should be dismantled and returned to Australia for display before it is blown away. But the Heritage Commission feels that it is the hut's location that gives it its significance. Australia's Antarctic Division hopes it can come to the rescue with a protective coating for the hut.

Charles Morgan

Japan's funding for unusual projects

Tokyo

QUANTUM electronic devices, pico-second chemistry, and 'plant ecochemicals' feature in the latest three additions to Japan's ERATO (Exploratory Research for Advanced Technology) programme.

The new projects, described at a symposium in Tokyo last week, will run for five years with budgets of ¥1,500 million–¥2,000 million (\$12–\$15 million) each.

ERATO, set up by the Science and Technology Agency's Research and Development Corporation of Japan in 1981, supports unusual, even eccentric, research that is too risky for industry but which may find applications in the long run. Each project is headed by a senior scientist who supervises a team of 15 to 20 young (under 35) researchers drawn from the universities, industry and government institutes. Foreigners are usually included to add an 'international' flavour.

Hiroyuki Sakaki of the Research Center for Advanced Science and Technology at the University of Tokyo heads the 'Quantum Wave' project which will investigate quantum effects in ultramicro-semiconductor devices. When atomic-sized structures (layers, boxes, spheres, wires) are constructed in semiconductors, the electrical and optical properties of the semiconductor can be drastically altered because of quantum wave effects that result from the imposition of additional structure at the atomic level. By harnessing and controlling these effects, Sakaki hopes to open up a new field of research.

The 'Micro Photoconversion' project, headed by Hiroshi Masuhara of Kyoto Institute of Technology, will use lasers to drive, measure and control chemical reactions at pico-second timescales in micrometer-sized reaction volumes. Practical applications are foreseen in electronics and new materials.

Junya Mizutani's 'Plant Ecochemical' project carries the most unusual title. Mizutani, who is in the department of agriculture at Hokkaido University, notes that whereas animals can run away when confronted with pests, plants have no option but to remain rooted to the ground and fight a chemical battle.

For example, the leaves of the walnut tree contain a derivative of an aromatic compound (1,4, 5-trihydroxynaphthalene) which is converted to the toxin juglone when the leaves fall to the ground. Juglone kills grasses at the base of the tree and also inhibits feeding by some beetles. Mizutani's team will look for 'ecochemicals' that may have application in agriculture.

David Swinbanks

Responsibility of scientists

SIR—Jamie Love (*Nature* 335, 758; 1988) asks that the scientific community organize to find constructive development in science, rather than wasting time and money in armament research.

Fortunately, some steps have already been taken. Pugwash is an organization of scientists that has been working for peaceful development since 1957. Recently, the Global Challenge Network was set up by Professor H.-P. Dürr (Munich). In 1986, the first international scientists' peace meeting was held in Hamburg, where the 'Hamburg proposals for disarmament' were announced. This year, the second international scientists' meeting 'From nuclear threat to mutual security' is being held from 2 to 4 December at Imperial College, London.

I should also mention national organizations such as the Union of Concerned Scientists (UCS) in the United States, Scientists against Nuclear Arms (SANA) in England and the natural scientists' organization Verantwortung für den Frieden (Responsibility for Peace), founded by West German scientists last spring.

All these activities are based on the

conviction that we, as scientists, should no longer let ourselves be used to develop means of mass destruction, no matter whether these are nuclear, chemical or biological weapons.

ULRIKE BEISIEGEL

*Universitäts-Krankenhaus Eppendorf,
Martinistr. 52,
2000 Hamburg 20, FRG*

SIR—Jamie Love seems not to have heard of Clausewitz's famed and sensible comment about war being a continuation of diplomacy by other means. As a result he has ended up putting the chariot before the horse.

War and the preparation for war cannot exist in the absence of some form of political or ideological conflict. It is only in the context of this conflict that armed forces are justified. These armed forces are required to meet the threat posed by the perceived enemy. Advantage is always sought and this leads, in the style of the Red Queen's race, to the development of newer weapons. The military spending that does this does not take place in a vacuum but only with the approval of a civilian government. It is the civilian gov-

ernment that defines the political stance of the nation and so defines any conflicts that the nation may have. Trying to stop the development of new weapons, or to slow the arms race, by appealing to those who design the weapons to stop it, does not address the underlying problems.

Disarmament in the absence of political resolution of a conflict has never prevented war. The best that can be said for it is that it has made subjugation less bloody. In the case of the Second World War, the failure to re-arm in good time (a logical equivalent of active disarmament) actually accelerated the outbreak of war. The German general staff had not planned to begin hostilities until some time around 1942–43. Similarly, technically superior artillery allowed the Prussians to go to war and win short, sharp victories against the Austro-Hungarian Empire and France. On the other hand, the Japanese managed to eliminate firearms from their culture in the seventeenth century. It did not stop them fighting, it just made the killing a little more difficult. The Intermediate Nuclear Force Treaty shows a measure of disarmament as a result of a political agreement. Even so, it looks a lot better than it is.

When there are no more unresolved territorial claims, when the mistakes of our forefathers are put right, when the wronged have justice, when there are no more have-nots, when there are no more oppressed people, then the underlying problems will be solved. Only when such a world has been built will we be able to stop building weapons. This is the great challenge. There is no place in it for rhetorical short cuts.

J. A. A. CHAMBERS
*6154 Cedar Crest Drive, Apt 107,
Johnston, Iowa 50131, USA*

SIR—Jamie Love does a great service in once more highlighting the dependence on scientists of those who wage war. Unfortunately, so long as military institutions exist, one can be sure that scientists will work for them, and their dismantling is a matter for other scientists only as individual parts of broader political movements. However, the penetration of the military into ostensibly academic institutions is also extensive and represents to many an unwelcome presence which sours their whole working environment. A useful exercise in scientific consciousness-raising and moral assertiveness could be performed if enough concerned scientists were willing publicly to forgo funds from military sources, however, innocuous sounding, and to reject collaboration with those who do. If this could evolve into a high-profile mutually supportive network, so much the better.

MICHAEL CLAGUE
*Section on Membrane
Structure & Function,
National Cancer Institute,
Bethesda, Maryland 20892, USA*

Disciplinary matter

SIR—J. Philippe Rushton (*Nature* 335, 8; 1988) may be right in his assertion that there is "widespread belief that there is no genetic basis by which people differ in important ways", but how that misconception results in the conclusion that "discipline and punishment are counterproductive" is at best a *non sequitur* and more likely a contradiction. If sociobiology is to live up to its own promissory notes, it will have to expunge infiltrators from the Dark Side.

JOHN HARTUNG

*Department of Anesthesiology,
State University of New York Health
Science Center at Brooklyn,
450 Clarkson Avenue,
Brooklyn, New York 11203–2098, USA*

Confusion in the lab

SIR—Those involved in running laboratories or working in them should be aware of the problems that could arise if the British Standards Institution (BSI) presses on with its revised standard for fume cupboards. Since 1982, they have been graded by a containment test, but BSI has now reverted to a test based on velocity.

The containment method provides better operator safety, and also provides scope for design development that may lead to technical innovation. At the same time, regulations for users due to be pub-

lished next year (Control of Substances Hazardous to Health Regulations 198–) for users call for tests of containment ability.

West Germany has a new standard based on the 1982 BSI draft, and in the Netherlands, after an investigation by the Dutch Institute for Environmental Hygiene and Health, there is now a specification for a 'standard product' based on containment rather than velocity.

Other manufacturers as well as users share my disappointment at the revised draft — and should we not in any case be aiming at a common standard with our European partners?

TONY HOARE

*Hoare Laboratory Engineering Ltd,
3 Victoria Road, Sevenoaks,
Kent TN13 1YB, UK*

Jenner request

SIR—I am doing research for a new biography of Edward Jenner based in part on papers held by the Library of the Wellcome Institute for the History of Medicine. Not only are these letters and diaries unused by previous twentieth-century biographers, but they were unknown to Jenner's original biographer, his student and friend, John Baron. I will be grateful for any advice from or documents held by Jenner descendants or enthusiasts.

RICHARD B. FISHER

*The Flat,
16 Highbury Place,
London N5 1QP, UK*

More ways than one for disorder

Despite half a century's hard work, there remain intolerably few circumstances in which the thermodynamics of disordered states can be calculated from plausible models.

With every month that passes, it becomes more and more apparent that the transition between microscopically ordered and disordered states is of enormous technical as well as theoretical importance, yet people's capacity to handle problems in this field seems merely to inch forward. Collective ingenuity, never at rest, meanwhile continues to invent new problems for solution. The common but depressing escape seems to be the devising of ever more intricate numerical simulations of particular problems, no doubt with the best intentions that these will be suggestive of more general understanding. But the hope is most often frustrated.

To be fair to the field and its practitioners, the most obvious problems are those which are for good reasons the most difficult to solve. Take any phase transition, say between a non-polar liquid, such as butane, and its vapour. One might assume that the vapour is a perfect gas, but the condensed (and more ordered) phase is an exceedingly disordered aggregation of molecules that must differ intrinsically from the perfect gas by virtue of the manner in which neighbouring molecules interact with each other. To say the least, they cannot be microscopic billiard balls. The mere fact that surfaces form between the liquid and the gas phase shows that two-body interactions are insufficient for an explanation. People hoping to make sense of this order-disorder transition must plunge into a theory of liquids, with its Kirkwood clusters and the like.

That is why much hope has always centred on the use of simpler systems as models of real-life transitions. Liquids and vapours differ from each other not merely in their degree of orderliness but also by the degree to which the free energy of each phase is dominated by the kinetic energy and, for the liquid, by the multi-molecular potential energy. For the study of order-disorder transitions, the ideal would obviously be a system in which there are two phases which are virtually identical except in their degree of order. What can such systems be?

Unhappily, the most obvious case, that of a ferromagnetic material, is far from easily handled. To be sure, the pre-conditions are readily satisfied. A ferromagnetic solid will seem much the same above and below the Curie temperature, so that experimental measurements, say of the specific heat, may be taken as direct measures of the energy involved in the transition from order to disorder. But

even here direct comparison between experiment and some theoretical model is far from simple. In real ferromagnets, it is not possible to make the simplifying assumption that only interactions between nearest neighbours matter, while even if it were, the resulting simplified model would prove to be intractable.

So much has been demonstrated by the large numbers of those who, in the past several decades, have broken their heads on just that task. The simplest model of a ferromagnet is a cubic lattice with little dipole magnets at each vertex. Even with the simplifying assumptions that the magnets can point in only one direction or its opposite, that only nearest-neighbour dipoles interact and that the lattice does not vibrate, whatever its temperature, it is not possible to find an analytical expression for, say, the free energy of this system. Even though the two-dimensional analogue of this problem was solved half a century ago by Lars Onsager, and while there are endless complicating modifications of it, the simple cubic-lattice model of a ferromagnet remains intractable.

That is why there must be great rejoicing that at least one class of problems, dealt with in the 1930s by people such as H.A. Bethe, is still firmly within the field of tractability, that of the solid solutions capable of forming ordered states. Brass, the alloy of copper and zinc, is one of these. Even with the correct proportions, the solid solution is a soft material which, when annealed, becomes hard and useable.

The thermodynamics textbooks give good accounts of the solution of such problems, which historically seem to have been the basis for the powerful way of treating these problems developed in the Soviet Union by the late Lev Landau and his then protégée, V. L. Ginsburg. Choose an "order parameter" (in the case of brass, the proportion of atoms out of place), assume the free energy to be some polynomial function of this parameter, calculate its minimum and then relate that to the physics of the problem.

This technique has been applied in fields as different as classical superconductivity and the superfluidity of helium. It works, even though there may be some trouble in telling just what physical significance to attach to the order parameter. But it would be better to start at the other end of the problem — with a model which is a plausible representation of a physical system — and to be able to calcu-

late the properties of the system from that.

The general frustration that these problems are usually intractable no doubt explains why so many other approaches to order-disorder problems have now been followed. One approach is to enumerate the possible configurations of a system numerically, by means of some mechanism dependent on the generation of random numbers. But the results remain obdurately numerical.

Much the same is true of attempts to calculate disordered structures such as the solid dendrites formed by crystallization from vapour or liquid. E. S. Langer was the first, two decades ago, to carry a macroscopic theory of a solidifying liquid onto a sufficiently microscopic scale to allow for processes which are limited by diffusion, but progress beyond that has been almost all numerical — E.T. Witten made a calculation, ten years ago, of diffusion-limited aggregation on a two-dimensional lattice that has since become the prototype for the numerical calculation of fractal structures formed under a variety of circumstances.

There is a neat version of such a calculation, backed up by experiment, in the recent literature (Helgesen, G. *et al. Phys. Rev. Lett.* **61**, 1736; 1988). Using plastic microspheres loaded with iron oxide, this group has been able to show both experimentally and by calculation how the fractal dimensions of the aggregates formed change systematically with the magnetic moment and also with the strength of an externally applied field.

But this is only scratching at the outside surface of an important set of problems. On the face of things, for example, what is called high- T_c superconductivity must crucially involve disorder of some kind. Departures from textbook stoichiometric compositions of the superconducting materials suggest disorder in the superconducting copper oxide layers. Similarly, the recent recognition of solid lattices with five- and ten-fold symmetry and the dependent issue of how these lattices grow by the aggregation of single atoms has provoked a series of questions intermediate between those raised by Witten and by the insoluble Ising lattice. And there is the whole string of problems involving macromolecules in gel formation, in diffusion through a gel and even in the transition from natural to denatured forms. It will be intolerable if all these problems remain only partly soluble for much longer.

John Maddox

Ozone depletion

Dynamics weaken the polar hole

Mark R. Schoeberl

JUST one year after the most spectacular Antarctic ozone depletion ever seen (October–November 1987; Fig. 1), the Nimbus-7 total-ozone mapping spectrometer (TOMS) has observed the weakest ozone hole since 1983. In 1987 the minimum value for the ozone column over Antarctica declined by 45 per cent during August and September from about 200 Dobson units (DU) to 110. In contrast, the 1988 minimum value in the ozone column fell from 210 to 180 DU over 20 days in early September, then rose slowly through most of the rest of September and October (Fig. 2). The polar vortex, a band of strong east–west stratospheric winds near 60° S, then broke up at the very end of October 1988, the earliest break-up observed in the past 10 years. The net spring total ozone decline for 1988 was less than 15 per cent. It is thus evident that although the chemistry of ozone depletion is understood, as confirmed by the *in situ* observations reported by J. W. Barrett *et al.* on page 455 of this issue¹ (and other observations to be published soon), changes in the physical dynamics of the polar atmosphere from year to year can have a powerful effect on the scale of the ozone hole.

The catalytic cycle implicated in the destruction of ozone relies on the release of active chlorine in the polar stratosphere as the Sun rises in the austral spring (see the recent News and Views articles by B. Thrush² and J. Pyle³). But for the cycle to produce the ozone hole, two key physical requirements must be satisfied.

First, air within the Antarctic stratosphere must be isolated from the mid-latitude environment, the polar vortex in effect forming a chemical containment

vessel⁴. Second, extremely cold temperatures must exist within the vortex (which can extend to about 60° S) to allow the formation of polar stratospheric clouds (PSCs). These provide surfaces for the heterogeneous reactions⁵ which release chlorine (Cl_2) into the atmosphere from the reservoir species HCl and ClONO_2 , while simultaneously removing odd nitrogen in the form of nitric acid (HNO_3), which remains with the PSC. Odd-nitrogen removal is an essential step which keeps active chlorine from returning to the reservoir compound, ClONO_2 . This removal can be further augmented by the condensation of the nitric acid trihydrate^{6–8} and the heterogeneous conversion of N_2O_5 and water ice to nitric acid⁹.

In the early-spring sunlight, the accumulated chlorine gas photolyses and combines with ozone to form chlorine monoxide (ClO) and oxygen. Ozone is destroyed catalytically with the formation of the chlorine monoxide dimer (Cl_2O_2) and its subsequent decomposition¹⁰ into $\text{Cl} + \text{Cl} + \text{O}_2$. Bromine may also play a role in enhancing the catalytic return of ClO to Cl through $\text{BrO} + \text{ClO} \rightarrow \text{BrCl} + \text{O}_2$.

Because the main catalytic path depends on the minor constituent chlorine monoxide reacting with itself, ClO mixing ratios of the order of 1 part per 10⁹ by volume are found to be necessary to explain the observed change in Antarctic ozone. The ground-based observations of Barrett *et al.*¹ and also the Airborne Antarctic-Ozone Experiment ER-2 observations of J. Anderson¹¹ show that this requirement was met during September 1987.

How cold and how isolated must the vortex be to produce an ozone depletion? Current observations of PSCs suggest that

the nitric acid trihydrate crystals (PSC type I) form at around 195 K with ice crystals (PSC type II) appearing near 187 K. Clearly warmer stratospheric temperatures will reduce the development of the PSCs, but the relationship between PSC distribution and the ozone depletion is probably not direct because air tends to move through PSCs. In other words, PSCs may act like flow-through chemical reactors, processing an air volume many times their observed spatial volume.

Competing with PSC production of chlorine is the leakage of the polar vortex containment vessel. Air crossing into the vortex will bring in not only ozone but also

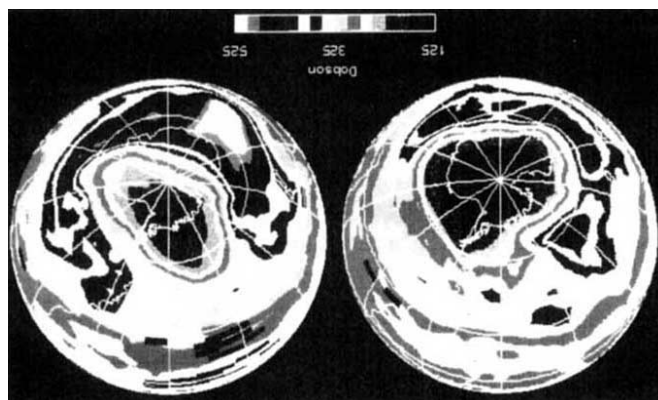


Fig. 2 Orthographic projection of the ozone distribution for 5 October 1987 (left) and 1988 (right). 5 October 1987 was the date of the lowest total ozone ever recorded by TOMS.

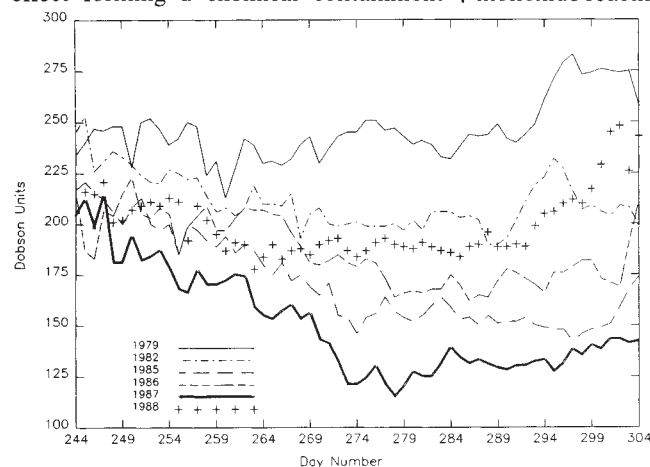


Fig. 1 The minimum total-ozone value observed by TOMS south of 30° S for September–October in 1979–1988. The data have been normalized to the Dobson network to correct for satellite instrument drift. Following the record minimum last year, the weaker ozone depletion this year perhaps indicates the action of dynamic effects.

odd nitrogen which rapidly reduces the amount of active chlorine. The ER-2 observations showed that the inner and outer regions of the containment vessel could be simply defined by the steep increase in ClO detected as the ER-2 flew north to south through the edge of the polar vortex. D. Hartmann¹² argues that air within the vortex was very tightly contained. He examined the variation of chemically long-lived species sampled during August and September 1987 and, after correcting for the adiabatic motion of the air parcels, found that their mixing ratios remained virtually unchanged throughout this period. This places a very small upper limit on the net transport of air across the vortex boundary during that period. A. Tuck¹³ argues, however, that at least the lower levels of the containment vessel are leaking, because the ER-2 found patches of extremely dehydrated air well outside the vortex.

From the middle of the 1988 austral winter, TOMS observed a rapid increase in sub-polar total-ozone variability. The change in total ozone coincided with the development of a warm stationary ridge in the lower stratosphere south of Australia, and a transition in winds at 20 km over Singapore from westerly to easterly. The expansion and contraction of the warm ridge frequently displaced the total-ozone minimum off the pole towards the Antarctic Peninsula. September satellite aerosol

observations showed only low-level PSCs within the polar vortex (P. McCormick, personal communication), agreeing with National Meteorological Center satellite retrievals which show that the polar stratosphere at 20 km was 10–15 °C warmer in September 1988 than in September 1987.

The increase in dynamical activity in the sub-polar stratosphere is undoubtedly responsible for the weakened ozone depletion this year, but according to the chlorine chemical theory at least two mechanisms may be involved. First, the presence of large-scale eddy activity in late winter increases the polar stratospheric temperature by increasing the adiabatic descent of air over the pole. This reduces the probability of PSC formation and therefore reduces chlorine production and odd-nitrogen removal. Second, a high level of sub-polar dynamical activity means that mid-latitude air is frequently entrained into the vortex, bringing in

more odd nitrogen. Again the result is a reduction in active chlorine. The important point is that both of these mechanisms act towards the same end, a reduction in the level of active chlorine. Thus the weak 1988 Antarctic ozone hole shows how sensitive this phenomenon is to year-to-year changes in dynamical activity despite the steady increase in stratospheric halocarbons. □

1. Barrett, J. W. *et al.* *Nature* **334**, 455–458 (1988).
2. Thrush, B. *Nature* **332**, 784 (1988).
3. Pyle, J. *Nature* **334**, 297 (1988).
4. Jukes, M. N. & McIntyre, M. E. *Nature* **328**, 590 (1987).
5. Solomon, S. *et al.* *Nature* **321**, 755–758 (1986).
6. Toon, O. B. *et al.* *Geophys. Res. Lett.* **13**, 1284 (1986).
7. McElroy, M. B. *et al.* *Geophys. Res. Lett.* **13**, 1296 (1986).
8. Crutzen, P. J. & Arnold, F. *Nature* **324**, 651–655 (1986).
9. Tolbert, M. A. *et al.* *Science* **240**, 1018–1021 (1988).
10. Molina, L. T. & Molina, M. J. *J. phys. Chem.* **91**, 433 (1987).
11. Anderson, J. G. *et al.* *NASA Conf. Pub.* **10014**, 143 (1988).
12. Hartmann, D. *NASA Conf. Pub.* **10014**, 169 (1988).
13. Tuck, A. *et al.* *NASA Conf. Pub.* **10014**, 165 (1988).

Mark R. Schoeberl is at the NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771, USA.

Gaia hypothesis

Can plankton control climate?

Tony Slingo

ONE of the more unusual contributions to the climate-change debate last year was the suggestion¹ that oceanic plankton exert an important, and perhaps controlling, influence on climate. The mechanism relies on the indirect effect, via dimethylsulphide (DMS) emissions, that plankton could have on cloud reflectivity. The authors, Charlson, Lovelock *et al.*, proposed that the feedback involved could contribute to Lovelock's Gaia hypothesis, by which the biosphere can maintain the climate within tolerable limits^{2,3}. On page 441 of this issue⁴, Schwartz reports that a search for similar climate effects from man-made sulphur dioxide reveals no such influence. Because sulphur dioxide should operate in the atmosphere in a similar way to DMS, this null result undermines the effect proposed by Charlson *et al.*, which is the only testable version of the Gaia hypothesis so far put forward.

In proposing their mechanism¹, Charlson *et al.* noted that plankton excrete DMS which is liberated into the atmospheric boundary layer to become oxidized to form sub-micrometre sulphate particles. These constitute the main source of the hygroscopic nuclei on which cloud drops form. Because the concentration of such cloud-condensation nuclei over the remote oceans is much lower than over land, changes in their number can affect the number of drops present in a cloud for a given liquid water content, and hence the mean size of the drops. This in turn alters the amount of solar radiation reflected back to space by the clouds and thereby affects the climate, because it is the

absorption of solar radiation which provides the energy input to the climate system. Charlson *et al.* argued that this chain represents a biogeophysical feedback between plankton and climate, because changes in climate would influence plankton population, which in turn would modulate the climate change, and so on.

Schwartz points out in this issue⁴ that global man-made emissions of sulphur dioxide are twice those from marine plankton. If Charlson *et al.* are correct, this suggests that anthropogenic emissions are an even more potent factor in climate change. Furthermore, because the anthropogenic sources are predominantly in the Northern Hemisphere and have appeared only during the past century or so, one would expect to see brighter clouds in the Northern Hemisphere (because of the enhanced nucleation and thus smaller drops) as well as a cooling relative to the less polluted Southern Hemisphere, over this period. Schwartz demonstrates that neither of these predictions are borne out by the data he has examined. He concludes that man has not significantly affected climate by his sulphur emissions, which means that the effect of plankton must be even smaller.

The satellite albedo data used by Schwartz certainly show no evidence for an enhanced contribution from clouds to Northern Hemisphere albedos. However, his cloud component of the total albedo (which is essentially the same as the short-wave cloud radiative forcing diagnostic introduced by Ramanathan⁵) is determined not only by the cloud brightness but also

by the cloud amount. The brightness is influenced by the column-integrated liquid-water content as much as by the mean drop size. Differences between cloud amounts and liquid-water contents in the two hemispheres could easily mask the effect of any differences between the mean drop sizes, giving a misleading signal in the overall albedo. By the same token, changes in these and other quantities over the period of the temperature record could have combined to minimize the inter-hemispheric differences in cloud forcing and climate change.

Despite these reservations, Schwartz's paper is a notable attempt to subject the DMS mechanism to observational test. Given the importance of cloud feedbacks in model estimates of climate change^{6,7} and the central role of cloud forcing to the mechanism, there is an urgent need to elucidate the factors which control both the amount and radiative properties of clouds. The current Earth Radiation Budget Experiment and International Satellite Cloud Climatology Project are thus timely, providing information on the radiation budget and cloud forcing⁸ and on cloud amounts and radiative properties⁹, respectively. But information on cloud liquid-water contents and mean drop sizes are also required. Techniques to estimate these from satellite data are available and should be applied to complete the picture. More observations are also needed (probably from airborne instruments) to establish more clearly the link between cloud condensation nuclei and mean drop size.

Many scientists are understandably reluctant to take the Gaia hypothesis seriously, and by association the mechanism proposed by Charlson *et al.*. This is unfortunate, because these authors have raised important questions regarding our understanding of the climate system. As with nuclear winter, it is not necessary to believe in a hypothesis to be stimulated by it to tackle *bona fide* problems which may have been neglected. The stimulus of nuclear winter led to practical work on the treatment of aerosol transport, scavenging and radiative transfer in numerical models, as well as giving insights into the importance of radiative-convective coupling. It is to be hoped that the papers by Charlson *et al.* and by Schwartz will provide a similar service in provoking work on the links between biological systems, atmospheric chemistry, cloud physics and climate. □

1. Charlson, R. *et al.* *Nature* **326**, 655–661 (1987).
2. Lindley, D. *Nature* **332**, 483–484 (1988).
3. Kerr, R. A. *Science* **240**, 393–395 (1988).
4. Schwartz, S. E. *Nature* **336**, 441–445 (1988).
5. Ramanathan, V. *J. geophys. Res.* **92**, 4075–4095 (1987).
6. Wetherald, R. T. & Manabe, S. *J. atmos. Sci.* **45**, 1397 (1988).
7. Schlesinger, M. E. *Nature* **335**, 303–304 (1988).
8. Harrison, E. F. *et al.* *Bull. Am. met. Soc.* **69**, 1144 (1988).
9. Schiffer, R. A. & Rossow, W. B. *Bull. Am. met. Soc.* **66**, 1498–1505 (1985).

Tony Slingo is in the Climate and Global Dynamics Division of the National Center for Atmospheric Research, PO Box 3000, Boulder, Colorado 80307, USA.

Protein structure

Method in MADness

Keith Moffat

A MAIN factor limiting the speed and accuracy with which protein structures can be determined by X-ray crystallography has been preparation of the series of crystals necessary for phase determination by the method of isomorphous replacement. Members of a series are intended to differ chemically and structurally in a very simple way from the original, native crystal, but seldom do; strict isomorphism is hard to achieve. The technique of multiple wavelength anomalous dispersion (MAD), relying on changes in the X-ray diffraction patterns from a single crystal arising from physical differences, rather than from many crystals arising from chemical and structural differences, has now been shown^{1,2} to yield excellent X-ray phases. Taking two proteins whose structure was previously unknown, a blue copper protein from cucumbers¹ and the biotin-binding core of streptavidin complexed with selenobiotin², Guss *et al.*¹ and Hendrickson *et al.*² have used the technique to determine the structures of these proteins at atomic resolution. The results suggest that the technique could be the long-sought general X-ray phasing technique for proteins and oligonucleotides.

The scattering of X-rays by free electrons in atoms and molecules can be perturbed by resonance between the electronic vibrations excited by the incident X-rays and the natural oscillations of the bound electrons, an effect termed anomalous (or resonance) scattering. Anomalous scattering occurs at all X-ray energies and wavelengths, and its presence makes the atomic scattering factor, f , complex: $f = f' + \Delta f' + i\Delta f''$; where f' is the atomic scattering factor in the complete absence of anomalous scattering effects, $\Delta f'$ is the (in-phase) dispersive term, and $\Delta f''$ is the (out-of-phase) absorptive term. The key to the MAD effect in X-ray phase determination lies in the fact that both $\Delta f'$ and $\Delta f''$ depend on wavelength, and their magnitudes are greatly enhanced³ when this wavelength approaches the atomic absorption edges of elements in the crystal. That is, the scattering contrast between the anomalous scattering atoms and the remainder of the structure is dependent on the X-ray wavelength, and is experimentally adjustable by tuning the X-ray source. In practice, the effect is restricted to the K-absorption edges of transition metals such as iron, copper and nickel that are intrinsic to certain protein crystals, or to the K- or L-edges of elements like platinum, mercury, selenium or the

lanthanides that can be incorporated into these crystals as extrinsic heavy atoms.

Anomalous scattering produces two effects on the observed X-ray intensities. First, the intensities at a single wavelength of the Bijvoet pairs for the reflections $\mathbf{k}$ and $\bar{\mathbf{k}}$ become unequal, and represent an anomalous or absorptive difference; and second, the intensities of a given reflection $\mathbf{k}$ (or $\bar{\mathbf{k}}$) differ with wavelength and represent a dispersive difference. Even when the effect is maximized, the changes in intensities are small: the scattering from metals represents only a small part of the total scattering from all atoms, and the anomalous component of the metals is even smaller. The r.m.s. values of both the anomalous and dispersive intensity differences constitute only 1–15 per cent of the total intensity, with lower values more common³. Highly accurate measurement of small intensity differences is essential.

Although it has been realized for more than 30 years⁴ that the MAD technique would allow direct phase determination, its successful reduction to practice has depended on the interplay between much more recent developments in technology and theory as well as exceptionally careful experimental design. First, the use of an intense, fully tunable synchrotron X-ray source permits the X-ray wavelengths to be carefully chosen, at the maximum of $\Delta f''$, the minimum of $\Delta f'$, and at two remote wavelengths, to maximize the signal and the phasing power in a way that is

impossible with laboratory sources. Second, multi-wire proportional chamber (MWPC) detectors⁶, originally designed for high-energy physics experiments, now provide high-precision recording of many reflections nearly simultaneously. Third, theory has been developed⁷ and reduced to practice⁸ that permits the 'best' phases to be determined from a set of MAD data. Fourth, and perhaps most important, great care has been taken^{1,2} to minimize systematic errors that plague accurate intensity measurements.

Crystals are carefully mounted to exploit their symmetry and morphology to minimize absorption errors arising from different X-ray path lengths for the reflections $\mathbf{k}$ and $\bar{\mathbf{k}}$; pleochroism⁹ arising from the local chemical environment of the anomalous scatterer that produces anisotropy in the anomalous scattering is considered; measurements at different wavelengths are interleaved to minimize the effects of drift in the source, X-ray optics and detector, and of radiation damage; and procedures⁸ for the local scaling of adjacent reflections are based on those originally developed for highly precise, single-wavelength measurements¹⁰. These eliminate slowly varying systematic errors, and identify and reject outlying observations that otherwise would have a far-reaching effect on efforts first to locate the anomalous scattering atoms, and then to use these locations for phase determination. In the MAD technique, all data are taken on a single crystal, eliminating the substantial inter-crystal absorption and scaling errors that limit the accuracy of the isomorphous replacement approach.

The proteins studied by Guss *et al.*¹ and by Hendrickson *et al.*² provide the most stringent test yet of the MAD technique as

Protein structures solved by the MAD technique

Protein	Anomalous scatterer*	Resolution (Å)	Synchrotron source [†]	Related structure known?	Ref.
Parvalbumin (<i>Opsanus tau</i>)	Tb, E	2.3	LURE	Yes	11
Cytochrome <i>c'</i> (<i>Rhodospirillum rubrum</i>)	Fe, I	6.0	PF	Yes	12
Azurin (<i>Pseudomonas denitrificans</i>)	Cu, I	3.0	CHESS	Yes	13
Haemoglobin (lamprey)	Fe, I	3.0	SSRL	Yes	14
Ferredoxin (<i>Clostridium acidi-urici</i>)	Fe, I	5.0	SSRL	Yes	8
Blue copper protein (cucumber)	Cu, I	3.0	SSRL	No [‡]	1
Core streptavidin–selenobiotin	Se, E	3.1	PF, SSRL	No	2

* I, Intrinsic heavy atom; E, extrinsic.

[†] LURE, Laboratoire pour l'Utilisation du Rayonnement Electromagnetique, Orsay; PF, Photon Factory, Tsukuba; CHESS, Cornell High Energy Synchrotron Source, Ithaca; SSRL, Stanford Synchrotron Radiation Laboratory, California.

[‡] This protein is a member of the class of blue copper proteins that includes plastocyanin and azurin, but differs from them to some extent in tertiary structure (see Fig. 5 of ref. 1).

the structure of the proteins involved was previously completely unknown. Elucidation of the crystal structure thus requires the best X-ray phases. Earlier studies^{8,11-14} examined either known structures, or new structures known to be closely related by sequence and functional homology to known structures (see table). Even though the known structures may not have been used directly in the MAD structure determination, subjective bias in the interpretation of the electron-density map of initially modest quality is impossible to exclude. As crystallographers know only too well, it is much easier to locate a known structural motif, for example a cytochrome fold, a calcium-binding EF-hand or a β -barrel, in such a map, than to identify a completely unknown motif.

Few proteins of interest to structural biologists are so obliging as to have a convenient, intrinsic heavy atom that exhibits anomalous scattering, so how is the technique to be generalized? First, crystallographers can now attempt to crystallize only a single heavy atom derivative in which a suitable extrinsic anomalous scattering atom has been chemically introduced, and ignore (at least initially) the native protein; the extent of isomorphism between the two is irrelevant to a successful structure determination via the MAD technique. More ingeniously, Hendrickson^{15,16} has proposed that selenomethionine may be biosynthetically incorporated into proteins in place of methionine. A methionine auxotroph of *Escherichia coli* grows quite happily on selenomethionine¹⁷, and can be used as the host for a recombinant plasmid designed to express the desired protein, in a selenomethionine medium^{15,16}. Because on average one amino-acid residue in 58 is methionine, the phasing power of such selenium derivatives is high^{2,16}. The approach is fairly general, and works also with *Pseudomonas aeruginosa*¹⁸. Extension to oligonucleotide structure determination via incorporation of bromouridine is clear.

A factor significantly limiting applications of the MAD technique is synchrotron beam time. Such studies require at least one week of actual data collection at the synchrotron². It is hard to conduct MAD experiments with the extreme care demanded for success when the beam-

Bulgarian geneticist recognized at last

It was not until June of this year, 40 years after his death, that Professor Doncho Kostov received the Dimitrov award — the highest national honour in Bulgaria — for his contribution to both international and Bulgarian science. The award marks the end of the period of intellectual vacuum suffered in Bulgaria since Stalin's time, and the eventual acknowledgement of the contribution of this eminent geneticist by his own nation.

Kostov died of a heart attack on 9 August 1949 at the age of 53. His death was probably related to the forced introduction of lysenkoism to Bulgaria the year before. Kostov believed passionately in scientific truth. There had been attempts to introduce lysenkoism in Bulgaria as early as 1947, and as the pillar of genetic research in the country, Kostov was an openly declared opponent of the principles of lysenkoism imposed from Moscow. Kostov's assistant professor, Dimitar Shishkov, was assassinated on 24 May 1947, a national holiday, and his murderers were never found.

Kostov graduated from the Institute of Agriculture in Halle, Germany and obtained his doctorate in Germany. He worked at Harvard University with Edward Murray East before a spell in Leningrad and Moscow between 1932 and 1939. In the Soviet Union he collaborated with Nikolai Ivanovich Vavilov, Nobel laureate Hermann Muller, Georgii Karpechenko, Peter Lisitzin and others, all of whom were opponents of Lysenko.

During this time, Kostov and his collaborators published many significant papers both in the Soviet Union and in the West. Until 1939, he was a regular contributor to *Nature*. He was the first to describe the

andrigene haploids in 1929; he studied the mutagenic action of alkaloids in 1930; and he photographed the discoid structure of the chromosomes of the fruitfly *Drosophila melanogaster*, also in 1930. Kostov's monograph on haploids, published in Holland in 1941, was a pioneering paper.

Kostov's major work "Cytogenetics of the genus *Nicotiana*" was published in 1943, simultaneously in Bulgarian and in English. This paper provided important proof of the role of the chromosome in evolution. It was this paper more than any which stalinist scientists used to harass Kostov to his death.

The Soviet geneticists with whom Kostov published some of his work, for example N.S. Arutunova, A. Orlov and N. Sarana, were killed during the Lysenko period because of the liberty that they had taken in publishing in the West. But as it turned out, only two geneticists in Bulgaria actually died as a result of stalinist victimization — Shishkov and Kostov. But about 100 other followers of the chromosome theory of evolution were tormented, dismissed or demoted. Their work could not be published for decades and they could not obtain academic qualifications.

Several historians attribute the lengthy obscurity of Bulgarian science to the period of Stalin's influence. But this can hardly be true for the whole of the past 40 years. In Bulgaria, as in other countries, the advent of lysenkoism was a moment when ignorance was raised to nationwide dimensions. Iskren Azmanov

Iskren Azmanov is in the Department of History at the Bulgarian Academy of Science, 1113 Sofia, Capaev Str. 52/17, Bulgaria.

line allocation committee is breathing down your neck, the X-ray beam or the detector electronics are briskly drifting, parts of the complicated apparatus keep breaking, your crystals are not cooperating, and your funding agency is complaining that your local and foreign travel budgets are enormous and always overspent. Plaintive comments along these lines that have escaped the principal authors' diplomatic axe can be found in many of the papers I have cited here^{1,2,8,11-14}. The next generation of dedi-

cated synchrotron radiation sources, the ESRF at Grenoble and the APS at Argonne, may alleviate the capacity problem by 1995 — but with the push in protein engineering and structure-based drug design, can the field wait?

Perhaps storage phosphor/image plate detectors^{19,20} will replace the precise but complicated and slow MWPC detectors; rather than scanning a narrow bandpass monochromator to four discrete wavelengths, perhaps a wider bandpass multi-layer monochromator in conjunction with unusual focusing²¹ or Laue oscillation²² techniques will be more effective; and perhaps the reliability, stability and accessibility of existing dedicated synchrotron radiation sources will be improved by their enthusiastic staffs. The future of the MAD technique, and of structural biology in general, will continue to depend in substantial measure on technological questions as well as scientific ones. □

Keith Moffat is in the Section of Biochemistry, Molecular and Cell Biology at Cornell University, Ithaca, New York 14853, USA.

- Guss, J.M. *et al.* *Science* **241**, 806–811 (1988).
- Hendrickson, W.A. *et al.* *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Lye, R.C. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **77**, 5884–5888 (1980).
- Hendrickson, W.A. *et al.* *Meth. Enzym.* **115**, 41–55 (1985).
- Okaya, Y. & Pepinsky, R. *Phys. Rev.* **103**, 1645 (1956).
- Phizackerley, R.P., Cork, C.W. & Merritt, E.A. *Nucl. instr. Meth.* **A246**, 579–595 (1986).
- Karle, J. *Int. J. quant. Chem.* **7**, 357–367 (1980).
- Hendrickson, W.A. *et al.* *Proteins* **4**, 77–88 (1988).
- Templeton, D.H. & Templeton, L.K. *Acta Cryst.* **A41**, 133–142 (1985).
- Hendrickson, W.A. & Teeter, M.M. *Nature* **290**, 107–113 (1981).
- Kahn, R. *et al.* *FEBS Lett.* **179**, 133–137 (1985).
- Harada, S. *et al.* *J. appl. Cryst.* **19**, 448–452 (1986).

- Korszun, Z.R. *J. molec. Biol.* **196**, 413–419 (1987).
- Krishna Murthy, H.M. *et al.* *J. biol. Chem.* (in the press).
- Hendrickson, W.A. *Trans. Am. Cryst. Ass.* **21**, 11–21 (1985).
- Hendrickson, W.A. in *Crystallography in Molecular Biology* (eds Moras, D. *et al.*) 81–87 (Plenum, New York, 1987).
- Cowie, D.B. & Cohen, G.N. *Biochim. biophys. Acta* **26**, 252–261 (1957).
- Frank, P. *et al.* *J. biol. Chem.* **260**, 5518–5525 (1985).
- Amemiya, Y. *et al.* *Nucl. instr. Meth.* **A266**, 645–653 (1988).
- Whiting, B.R., Owen, J.F. & Rubin, B.H. *Nucl. instr. Meth.* **A266**, 628–635 (1988).
- Arndt, U.W. *et al.* *Nature* **298**, 835–838 (1982).
- Moffat, K., Szebenyi, D. & Bilderback, D. *Science* **223**, 1423–1425 (1984).

X-ray astronomy

Sources of variable oscillations

N.E. White

THREE years have passed since the discovery¹ of 20–50-hertz quasiperiodic oscillations (QPOs) from the bright X-ray source in the Galactic Centre, GX5–1. This discovery, made using the European Space Agency's Exosat satellite observatory, may provide the first direct indication that this class of X-ray binary contains weakly magnetized neutron stars with millisecond rotation periods. The considerable progress that has been made in classifying the QPO phenomenon and in developing detailed models was discussed at a recent meeting*. Quasiperiodic oscillations have now been detected not only from sources similar to GX5–1, but also X-ray pulsars and black-hole candidates. It also has reawakened interest in the time variability properties of X-ray sources in general. This is well illustrated by the paper from Miyamoto *et al.*, on page 450 of this issue², in which an analysis technique used for studying QPOs is applied to gain new insight into the time variability of Cyg X–1.

The quasiperiodic oscillations of GX5–1 were identified¹ during a search for millisecond pulsars. But rather than finding steady pulsations at a single frequency, the observers found a broadly peaked distribution of frequencies (recorded with a broad-band X-ray detector). The maximum frequency varied significantly with the long-term variations in average intensity (Fig. 1). Following the GX5–1 discovery, Exosat revealed a further eight sources that show QPOs with frequencies of 6–50 hertz. Amongst these are Sco X–1 and Cyg X–2, both of which lie away from the obscuring material in the Galactic Centre and are well known to optical observers. They have orbital periods of 0.7 and 9 days, respectively, each comprising a late-type giant star losing mass to a neutron star.

Many different physical processes can give rise to QPOs and many models were proposed to explain the GX5–1 result. The most appealing was the beat-frequency-modulated accretion model^{3,4}, which requires a neutron star with a 10^8 – 10^9 -gauss magnetic field interacting with an accretion disk. The compression of a magnetosphere as the accretion rate increases explains a strong QPO-frequency/intensity correlation found for GX5–1. The magnetic field strength and spin period predicted by this model are very similar to those found for the millisecond radio pulsars, of which systems like GX5–1 are thought to be the progenitors. But

the frequency/intensity correlation was soon found not to be a universal feature. In many it was absent and in some, Sco X–1 for example⁵, the correlation was sometimes present and sometimes not.

QPOs are now classified into two distinct modes⁶. In the slow, 6-hertz mode, the frequency shows little correlation with intensity, and in the fast, 20–50-hertz mode there is a strong frequency/intensity correlation. These two QPO modes correspond to different spectral states of the source⁷. This is best seen in a colour–colour diagram (Fig. 2). Each source follows a characteristic Z-shaped pattern with each branch of the Z showing only one particular QPO mode. The fast QPOs occur on the horizontal branch (HB) and

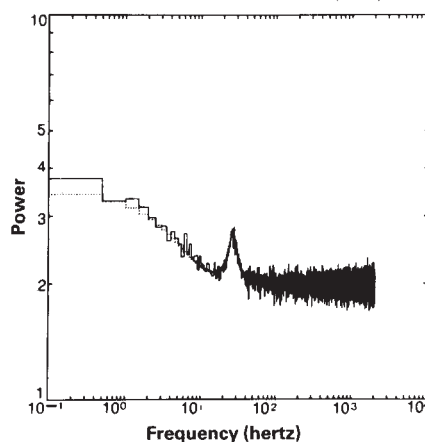


Fig. 1 Power spectrum for GX5–1. At low- and high-frequency extremes, the spectrum represents noise. The central peak corresponds to the quasiperiodic oscillations, varying around a frequency of 27.6 hertz. The peak frequency of the QPOs increases with increasing average power. (From ref. 1.)

the slow QPOs appear on the lower part of the normal branch (NB). For Sco X–1 only, as the source turns the corner to go from the normal branch to the flaring branch (FB), the 6-hertz QPOs begin to increase frequency and is again correlated with intensity. As with every classification scheme there are exceptions, such as the rapid burster (L. Stella, Exosat Observatory) and Cir X–1 (A.F. Tennant, Marshall Space Flight Center), but the above scheme seems to work well for most QPO sources (G. Hasinger, Max Planck Institute, Garching). Evidence was presented for a second class of bright Galactic Centre sources which follow a different pattern in the colour–colour diagram and do not show QPOs (M. van der Klis, Exosat; Hasinger).

The spectrum that underlies the Z-pattern is a combination of a black-body

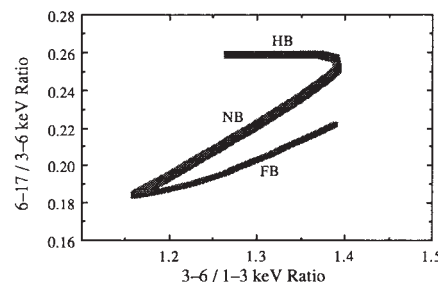


Fig. 2 An X-ray colour–colour diagram constructed by dividing the counts in three different energy bands. Three branches are visible: a horizontal branch (HB) where fast 20–50-hertz frequency/intensity-correlated QPOs occur, a normal branch (NB) where slow 6-hertz frequency/intensity-independent QPOs occur and the flaring branch (FB) where again 10–30-hertz frequency/intensity-correlated QPOs are seen (only from Sco X–1).

(thermal) emission attributed to the neutron star, and an exponential-like component, presumably from an accretion disk. On the horizontal and normal branches, the exponential components dominate. Dramatic increases in the black-body component cause the flaring branch. Comptonization, the degradation of X-rays by electron scattering, was discussed as a promising spectral formation mechanism and variations in optical depth may account for the normal and horizontal branches of the Z-pattern (F.K. Lamb, University of Illinois; G. Miller, University of Illinois; R. Wijers, University of Amsterdam; myself).

Two new models were proposed for the slow, 6-hertz mode. One involves 6-hertz oscillations in an extended atmosphere that blankets the neutron star at the 'Eddington' limit, for which outward radiation pressure balances accretion (N. Shibazaki, Rikkyo University). The other unifies the spectral and temporal properties (Lamb) and finds 6-hertz oscillations from an instability in the radial flow close to the neutron star, again only at the Eddington limit. The beat-frequency model is in both cases still invoked to explain the fast 20–50-hertz mode.

Scattering in a cloud surrounding the neutron star would smear out any millisecond pulsations from the neutron star rotation but leave the QPOs unaffected (N. Kylafis, University of Crete). The fast QPOs of high-energy X-rays from Cyg X–2 and GX5–1 lag behind the lower-energy ones by 4 and 0.4 milliseconds, respectively (van der Klis). These lags were first thought to be caused by Compton up-scattering (X-rays gaining energy from fast electrons). But such a lag, in detailed calculations, requires a cloud radius of about 1,000 kilometres (R.W. Bussard, University of Arizona), which is too large because other estimates show that most of the observed energy should be released within a few neutron-star radii (10–30 kilometres). Even worse, the slow QPOs from Cyg X–2 show a time lag of 70

*Quasiperiodic Oscillations in Luminous Galactic X-ray Sources, La Cienaga, New Mexico, 2–7 October 1988.

milliseconds, or almost half a QPO cycle (K. Mitsuda, Institute of Space and Astronautical Science, Tokyo), which requires even larger cloud sizes. The cloud size can be reduced somewhat by matching the temperature of the electrons and photons (Lamb), but the feeling was that, for the 6-hertz mode at least, the time lag must be intrinsic to the QPOs. Energy-dependent time lags in the rapid time variability of Cyg X-1, as Miyamoto *et al.* argue¹ in this issue, are also not consistent with the simple comptonization model and also must be generated in the same process as the variability.

The beat-frequency model should also work for X-ray pulsars and this idea prompted several searches, with two notable successes. The 42-second pulsar EXO2030+375 shows 0.2-hertz QPOs (L. Angelini, Exosat) and the 4.8-second pulsar Cen X-3, 0.03–0.07-hertz QPOs (Tennant).

The possibility was also discussed that some of the QPO modes might be oscillations excited in an accretion disk (H. van Horn, University of Rochester), although

this is at present no more than an idea. An interesting recent development is that observations from the satellite Ginga of several black-hole candidates have revealed QPOs (Mitsuda). The most notable is a strong 6-hertz oscillation detected from GX339-4 that is very similar to that seen from the GX5-1 sources on the normal branch. These QPOs appear only when there is a high-energy tail to the normally ultra-soft spectrum. The QPOs are largely confined to this tail. As yet only black-hole candidates identified on the basis of the similarity of their X-ray spectrum to that of Cyg X-1 have shown QPOs, so there remains some doubt as to whether or not these truly contain black holes. □

1. van der Klis, M. *et al.* *Nature* **316**, 225–230 (1985).
2. Miyamoto, S., Kitamoto, S., Mitsuda, K. & Dotani, T. *Nature* **336**, 450–452 (1988).
3. Alpar, M.A. & Shaham, J. *Nature* **316**, 239–241 (1985).
4. Lamb, F.K. *et al.* *Nature* **317**, 681–687 (1985).
5. Pridorsky, W. *et al.* *Astrophys. J.* **306**, L95 (1986).
6. Stella, L. in *Plasma Penetration into Magnetospheres* (eds Kyfalis, N., Papamastorakis, J. & Ventura, J.) 199–214 (Crete University Press, 1986).
7. Hasinger, G. *Astr. Astrophys.* **186**, 153–158 (1987).

N.E. White is at the Exosat Observatory, Space Science Department of ESA, 2200 AG Noordwijk, The Netherlands.

High-energy physics

Extreme states of nuclear matter

Helmut Satz

WHEN two heavy nuclei collide at high energy, they create a many-body system with a density much higher than normally found anywhere. In such collisions it should be possible to investigate the thermodynamics of strongly interacting matter — consisting of protons, neutrons and other hadrons, or of the quarks from which these are made up — and in doing so, study on a small scale the environment of the very early Universe. The theory of strong interactions, quantum chromodynamics (QCD), predicts that the hadrons should dissolve into a plasma of deconfined quarks if its energy density is sufficiently high. Several experiments at CERN, the European nuclear research centre, and Brookhaven National Laboratory were discussed at a recent meeting on ultra-relativistic nucleus–nucleus collisions*. The results show that such nuclear collisions are more than mere superpositions of many nucleon–nucleon collisions and there are even signs of the predicted new state of matter, the quark–gluon plasma.

To study the thermodynamics of strongly interacting matter, it is necessary to produce a 'bulk' system, containing many hadrons and of a size much larger than the range of strong interactions. Its constituents should be in thermal equilibrium, and its energy density should be much higher than that of normal nuclear matter for the regime of new phenomena to be reached.

Such matter could be characteristic of the state of the Universe in the first 10^{-5} seconds of its existence, and hence have very unusual properties.

To achieve such an energy density, ions in the mass (A) range from oxygen ($A = 16$) to sulphur ($A = 32$) were collided with stationary heavy targets (gold to uranium). CERN provides beam energies of up to 200 gigaelectronvolts (GeV) per nucleon, Brookhaven of 15 GeV per nucleon. With existing facilities, neither laboratory can accelerate ions heavier than those mentioned, but both plan extensions to accommodate all masses (N. Samios, Brookhaven; R. Stock, University of Frankfurt).

How large is the droplet of matter produced in high-energy nuclear collisions? This question was addressed in the CERN experiment¹ NA 35, which used the Hanbury-Brown–Twiss technique of astronomy, in which the interference of light from two incoherent source points allows a determination of the size of the emitting system. Applied to the interference of particles produced in nuclear collisions, this method gives information about the size of the interaction region as the nuclear fireball 'freezes out'. The results are very encouraging: whereas in a proton–proton collision at 200-GeV beam energy the typical emitting volume is about $5\text{--}10\text{ fm}^3$ ($1\text{ fm} = 10^{-15}\text{ m}$), the volume in the corresponding collision of

an oxygen beam with a gold target is about $1,000\text{ fm}^3$ — two orders of magnitude larger (J.W. Harris, Lawrence-Berkeley Laboratory). Heavier projectiles in future experiments could increase this value by at least another power of ten.

The volumes just discussed correspond to a comparatively late stage in the history of the bubble. Very shortly after the collision, we expect to find a hot system which quickly expands and thereby cools off; only then does it form and emit the observed hadrons^{2,3}. What were the highest energy densities attained earlier in the collision? This can be estimated by measuring the overall energy carried by the emitted particles and dividing this by the initial collision volume. Most experiments at CERN and Brookhaven performed measurements of this type. Because of its higher incident beam energy, CERN also attained higher energy densities with values reported up to $2\text{--}3\text{ GeV fm}^{-3}$. Such densities are $15\text{--}20$ times higher than those of normal nuclear matter (0.15 GeV fm^{-3}) and considerably exceed even the energy density inside a nucleon (0.5 GeV fm^{-3}). Extrapolating the present results to heavier beam nuclei suggests that values up to 5 GeV fm^{-3} can be reached. How high the energy density needs to be for the transition to the quark–gluon plasma has been a central topic in statistical QCD⁴. Lattice calculations (reviewed by A. Ukawa, CERN) predict that deconfinement should occur for energy densities of $1.5\text{--}2.5\text{ GeV fm}^{-3}$. The uncertainty arises mainly from the small lattice size in present evaluations; future calculations on larger lattices should make it more precise.

Do nuclear collisions lead to systems in thermal equilibrium? A much discussed result from one of the Brookhaven experiments (E 802) appears to shed some light on this problem⁵. The Japanese–American collaboration measured the kaon/pion production ratios K^+/π^+ and K^-/π^- for a 15-GeV-per-nucleon silicon beam incident on a gold target. At this energy, proton–proton collisions give $K^+/\pi^+ \approx 0.1$, $K^-/\pi^- \approx 0.05$; however in nuclear collisions, the K^+/π^+ ratio becomes much larger (about 0.25), whereas K^-/π^- remains roughly the same (P. Vincent, Brookhaven). At the Brookhaven beam energy, the baryon number density of the produced system is particularly high, because the two colliding nuclei almost stop each other within a small region of space.

A high baryon number density leads for an equilibrium system to an enhancement of positive kaons, because processes such as $\pi^+ + p \rightarrow \Lambda + K^+$ or $\pi^+ + n \rightarrow \Lambda + K^+$ reduce the number of nucleons by transforming them and thus reduce the nucleon Fermi pressure, which arises from the quantum rule preventing identical particles from occupying the same region of space.

*Quark Matter 1988, Lenox, Mass., 25–30 September 1988.

The processes also require less input energy than the production of a K^+K^- pair, which is the only way to make negative kaons. Different theoretical models, based on an equilibrium hadron gas or on an initial quark–gluon system with subsequent expansion and hadronization, lead to values in accord with those measured at Brookhaven (B. Friman, GSI Darmstadt). These data thus support the idea that the systems produced are in thermal equilibrium; a similar conclusion is drawn from the increase in Λ production observed at CERN. Moreover, such an equilibrium is reached more easily from an initial quark–gluon plasma than from a conventionally interacting hadron gas⁶.

Has the quark–gluon plasma been produced in the experiments carried out so far? If formed, it will rapidly expand, cool off and revert back to the hadronic phase. We are thus looking for evidence of a deconfined system which no longer exists as such when measurements are taken. T. Matsui and I predicted⁷ that a suppression of the production of J/ψ mesons, comprising a charmed–anticharmed quark pair ($c\bar{c}$), in nuclear collisions would provide a crucial hint in this search. First measurements made at CERN by the NA38 collaboration indeed indicate that such a suppression occurs⁸. These results have now been extended to sulphur as well as oxygen beams and are now based on much more data (J. Y. Grossiord, University of Lyon). Quark deconfinement results in a reduction of J/ψ formation because in a

deconfining matter the c and $\bar{c}$ constituents cannot bind to each other and hence eventually hadronize with other quarks as open charm mesons. The experiment, for both ^{16}O and ^{32}S beams, shows more than 50 per cent suppression at the highest energy densities attained; the observed momentum distributions are in accord with deconfinement predictions as well⁹.

Many theorists have sought a mechanism other than deconfinement that can give the observed effect (J.-P. Blaizot, Saclay). Conventional alternatives, such as absorption in normal nuclear matter, cannot reproduce the observed behaviour. But several groups have shown that an absorption of the J/ψ in very dense hadronic matter, when combined with initial-state gluon scattering, can account for the data (S. Gavin, Lawrence-Berkeley Laboratory). It appears clear now that the observed J/ψ behaviour is due to the formation of extremely dense and reasonably long-lived strongly interacting matter. Whether global 'colour' screening (quark deconfinement) or local scattering (absorption) is the decisive mechanism for suppression can be answered only by studies with heavier nuclei. □

1. Humanic, T. J. *et al.* *Z. Phys.* **C38**, 79–84 (1988).
2. Anisshetty, R. *et al.* *Phys. Rev.* **D22**, 2793–2804 (1980).
3. Bjorken, J. D. *Phys. Rev.* **D27**, 140–151 (1983).
4. Fukugita, M. *Nucl. Phys.* **B4**, 105–129 (1988).
5. Miake, Y. *et al.* *Z. Phys.* **C38**, 135–140 (1988).
6. Koch, P. *et al.* *Phys. Rep.* **142** 167–252 (1986).
7. Matsui, T. & Satz, H. *Phys. Lett.* **B178**, 416–422 (1986).
8. Bussière, A. *et al.* *Z. Phys.* **C38**, 117–128 (1988).
9. Karsch, F. & Petronzio, R. *Phys. Lett.* **B193**, 105 (1987).

Helmut Satz is in the Fakultät für Physik, Universität Bielefeld, D-4800 Bielefeld, FRG.

Immunology

Suppression of the immune response by cytotoxic T cells

Elizabeth Simpson

IMMUNE responses of higher animals have to discriminate between self and non-self, be specific and be regulated. The bone marrow produces B cells, which make antibodies in response to foreign antigens. T cells are produced by the thymus and one of their functions is to help the B cells to respond to antigen. Shinohara *et al.*, in their paper on page 481 of this issue¹, investigate the nature and mode of action of T lymphocytes involved in regulation of the immune response. They provide evidence that cytotoxic T cells, the T lymphocytes that kill infected or foreign cells, can also kill antigen-specific B cells, thus mediating the suppression of specific B cells in the immune response.

The ability to discriminate between self and non-self is mediated by selection of T cells as they differentiate in the thymus: those that react with high affinity to self-antigens are eliminated. The population of T cells that migrates into the blood and

lymph is depleted of the most autoreactive cells^{2,3} and is enriched for two types of T cells with particular affinity for major histocompatibility complex (MHC) molecules⁴. (T cells recognize foreign antigen only when the antigen is associated with MHC molecules.) One type consists of T cells carrying the CD8 cell-surface marker (CD8^+), and these bear T-cell receptors specific for foreign (non-self) peptide antigens bound to MHC class I or, occasionally, to MHC class II^{1,6} molecules. The other type consists of CD4^+ T cells with receptors specific for antigen plus class II molecules.

T cells with CD8^+ markers are cytotoxic and T cells with CD4^+ markers are termed 'helper' T cells. But it is known that CD4 cells can also kill (see ref. 1). It was believed that suppressor T cells, which suppress (down-regulate) immune responses, are a class apart, variously designated as a subclass of either CD8^+ or CD4^+ T cells,

depending on the experimental system. The failure to isolate clones of suppressor T cells, or biochemically to characterize their receptors or products, has generated scepticism — not with respect to the occurrence of down-regulation or suppression as a phenomenon, but as to whether there are T-suppressor cells separate and distinct from CD4^+ T-helper cells or CD8^+ cytotoxic T cells.

What Shinohara *et al.* now show¹ is that CD4^+ and CD8^+ T cells, specific for an antigenic epitope termed the 'carrier', can kill B cells bearing antibody molecules specific for an antigenic epitope termed the 'hapten' in the presence of the hapten–carrier molecular complex. The B cells bind the complex, internalize and process it, then present the carrier epitope in association with their own MHC class II molecules to the cytotoxic cells, which kill the B cells. Shinohara *et al.* show that the B-cell killing, or lysis, is class II-restricted. They do this by blocking antibody binding to both CD4^+ and CD8^+ carrier-specific T cells. There have been previous reports of lysis by class II-restricted CD4 cells but the suppressor function of these cells has been obscured in systems in which their helper function predominates. Lytic CD8^+ T cells specific for soluble antigens have been reported previously⁶, but these were class I-restricted: Shinohara *et al.* show that class II-restricted CD8^+ cells are even more powerful suppressors, as they appear to be lytic to B cells in the presence of very low (probably physiological) levels of antigen.

Lanzavecchia *et al.* showed recently⁷ that human T cells can process and re-present, in association with endogenous MHC class II, any antigenic molecule which could bind to a surface receptor such as CD4 or CD8 , and then be internalized, processed and re-presented in association with the class II molecules of the processing cell. Such T cells could then become the target for elimination by antigen-specific cytotoxic T cells. The data of Lanzavecchia *et al.*⁷ and of Shinohara *et al.*¹ suggest that this suppression can be mediated by CD4^+ and CD8^+ cytotoxic T cells which can eliminate all antigen-presenting cells. When such cells are professional antigen-presenting cells, or T or B lymphocytes specific for antigen, the likely result is down-regulation of the immune response. □

1. Shinohara, N., Watanabe, M., Sachs, D. H. & Hozumi, N. *Nature* **336**, 481–484 (1988).
2. Kappler, J. W., Staerz, U., White, J. & Marrack, P. C. *Nature* **332**, 33–40 (1988).
3. MacDonald, H. R. *et al.* *Nature* **332**, 40–45 (1988).
4. Kisielow, P., Blüthmann, H., Staerz, U. D., Steinmetz, M. & von Boehmer, H. *Nature* **333**, 742–745 (1988).
5. Teh, H. S. *et al.* *Nature* **335**, 229–233 (1988).
6. Staerz, U. D., Karasuyama, H. & Garner, A. M. *Nature* **329**, 449 (1987).
7. Lanzavecchia, A., Roosnek, E., Gregory, T., Berman, P. & Abrignani, S. *Nature* **334**, 530–532 (1988).

Elizabeth Simpson is in the Transplantation Biology Section of the Clinical Research Centre, Watford Road, Harrow, Middlesex HA1 3UJ, UK.

Gene regulation

A hit-and-run mechanism for transcriptional activation?

Walter Schaffner

At present it is a matter for debate whether enhancer/upstream promoter sequences of mammalian genes and the transcription factors binding to them are required for each round of transcription initiation, or whether they are needed only as a trigger to organize a gene into a stable transcription complex, after which transcription by RNA polymerase II can continue without these DNA sequences and their bound factors. Data compatible with a transient enhancer requirement were first provided¹ by Wabl and Burrows, who showed that immunoglobulin gene transcription persists in a B-cell line in which the enhancer had been removed by spontaneous deletion. The recent experiments of two groups^{2,3} on transcription from a viral promoter *in vitro* at first sight seem to suggest an explanation at the molecular level for these and subsequent observations *in vivo*. There are, however, alternative explanations for both sets of phenomena, and it may be necessary to look elsewhere for the molecular basis of cellular memory.

When Wabl and Burrows first discovered the apparently transient need for the immunoglobulin enhancer, they considered the enhancer to be an artefact of transfection. The remarkable tissue specificity of the enhancer, however, made this seem unlikely, and in collaboration with Andreas Radbruch and Sigi Klein, we undertook a more detailed investigation of a cell line that continued to produce immunoglobulin transcripts after spontaneous deletion of the immunoglobulin heavy-chain enhancer. One possible explanation was that the cell had mutated to render the presence of an enhancer dispensable for the efficient expression of the endogenous or a transfected immunoglobulin gene. Alternatively, it was also conceivable that the deletion had, by chance, created a substitute transcription factor binding site(s). But when we cloned the endogenous gene and reintroduced it into the same cell line (after appropriate tagging), it was expressed only when linked to an enhancer¹. At the same time, essentially the same gene, but devoid of the enhancer, was strongly transcribed in its genomic context. Thus, both the above explanations were excluded. We argued at the time that there are only two reasonable explanations for these findings: first, that there is another, as yet unidentified enhancer that substitutes for the deleted immunoglobulin heavy-chain enhancer; or second, that the enhancer organizes the DNA into a stable

transcription complex, after which it is no longer required.

It is this latter possibility that seems to be supported by two recent papers on the initiation of transcription from one of the promoters of adenovirus. For several years, Robert Roeder and his colleagues have been analysing the mechanism of transcriptional initiation by RNA polymerase II. Whereas their earlier work concentrated mainly on the adenovirus major late promoter, they have now, together with the group of Michael Green, looked at the E4 promoter of adenovirus.

This promoter is particularly suitable because transcription *in vitro* can be stimulated 1,000-fold by the enhancer/upstream binding factor ATF. This cellular transcription factor binds to multiple sites upstream of the E4 TATA box. In a series of experiments^{2,3}, Roeder, Green and collaborators show that ATF facilitates formation of an initiation complex by TFIID (TATA box factor), RNA polymerase II, TFIIB and TFIIIE. The formation of this complex is revealed by an extended footprint over the TATA box and downstream of it. Removal of ATF, using a large excess of competitor oligonucleotide with an ATF-binding site, does not eliminate the extended TATA box footprint, nor does it prevent subsequent transcription by RNA polymerase. The authors conclude that ATF is required only transiently for the process of transcription initiation. In other words, ATF is required for efficient assembly of the

Oldest known reptile found in Scotland

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

THE discovery in Scotland of the earliest known fossil reptile is the most important fossil find in the past 50 years, in the opinion of Timothy Smithson (University of Newcastle-upon-Tyne), who identified it. At 340 million years old, it is the oldest animal of its kind by 40 million years, and is contemporaneous with some of the most ancient faunas containing tetrapods of any kind (see *Nature* 333, 768; 1988). The 8-inch-long articulated skeleton was found in Lower Carboniferous rocks in southern Scotland by professional collector Stan Wood, who announced the find at the British Museum (Natural History) in London on 16 November. The fossil, seen in the figure, will be on display there until 17 January. The inset shows an artist's impression (by Michael Coates) of how the reptile might have looked.

Wood has found many spectacular Lower Carboniferous amphibians, but this is his first reptile, marked out as such by anatomical details of the skull table, vertebral column and feet. The discovery will prompt a radical revision of early amniote phylogeny, says Michael Benton (Queen's University, Belfast); that fully evolved reptiles existed in the Lower Carboniferous throws doubt on the supposedly close phylogenetic links between amniotes and anthracosaurs, a group of fossil amphibians.

Potassium-argon isotope dates of 338–340 million years old for the strata were established before the new fossil was discovered. Lower Carboniferous Scotland lay across the Equator, and the humid, tropical jungle was frequently inundated with volcanic ash. This earliest-known reptile may have been boiled to death in a hot spring and covered with dust, which could account for its excellent preservation. Henry Gee.

initiation complex but is dispensable thereafter.

On the face of it, these data could provide a molecular explanation for the data on immunoglobulin gene transcription *in vivo*, suggesting that an enhancer is needed only to establish transcription of a gene and is later dispensable. However, a closer look suggests a different interpretation. Removal of the bound ATF after formation of the initial complex reduces the level of transcription *in vitro* to about 50 per cent, which corresponds to the level observed if reinitiation, but not elongation, is blocked (for example, by the nonionic detergent sarkosyl). Thus the transcription observed after the removal of ATF could represent elongation of existing transcripts, ATF still being needed for initiation. Regarding the transient requirement for the immunoglobulin enhancer *in vivo*⁴, the alternative explanation is a second, undiscovered enhancer. This explanation seemed quite dull — it did not entail any new concepts, and in any case no-one was able to find this hypothetical enhancer within a dozen or so kilobases from the immunoglobulin gene boundaries. Nonetheless, it now turns out that it is probably correct.

With a series of well-designed experiments, Rudolf Grosschedl and Mary Marx have apparently settled this question⁵. They introduced an immunoglobulin μ -gene containing the enhancer flanked by D and J recombination signals into pre-B cells by stable transformation. They found that subsequent deletion of the enhancer from the transformed gene, by D- to J-joining, reproducibly leads to the elimination of μ -gene transcription. Grosschedl and Marx conclude that the enhancer does not confer upon the μ -gene a memory of its active state. Thus, the enhancer seems to be required both for establishment and for maintenance of μ -gene expression. Grosschedl and Marx also verify that the μ -gene is hypomethylated before transcription is switched off (and that subsequent deletion of the enhancer results in methylation of the gene). This is important because it has been argued by others⁶ that one of the effects of an enhancer is to help remove the methyl groups and to reorganize the DNA domain into active chromatin, after which it becomes dispensable for continued transcription. This would make it necessary to postulate not only the removal of an activating factor, but also some repressor mechanism for every inducible promoter to switch off transcription again.

The enhancer effect is exerted, however, at the level of transcription initiation. Weak and strong enhancers can be created by multimerization of factor binding sites, resulting in a stepwise increase in transcription rate. This suggests that transcriptional enhancement is not an all-or-none phenomenon⁷. In other

words, an RNA polymerase II transcription unit does not exist in only two states, 'on' or 'off', as seems to be the case for the ribosomal RNA transcription units⁸.

This difference between RNA polymerase II transcription units and ribosomal RNA genes, which are transcribed by RNA polymerase I, may reflect the fact that there are multiple copies of the ribosomal RNA genes whereas most genes transcribed by RNA polymerase II are single-copy. Thus, ribosomal RNA gene transcription can be modulated by regulating the number of transcription units in the 'on' state; and this may also be true for genes transcribed by RNA polymerase III⁹ and perhaps even for the multi-copy U small nuclear RNA genes¹⁰. By contrast, all-or-none regulation of transcription just would not make sense for most of the genes transcribed by RNA polymerase II, because it could severely limit the degrees of freedom of gene regulation. This should not, however, be taken to mean that irreversible mechanisms of gene activation/repression do not exist. A differentiated state, for example, can be triggered by a stimulus that does not have to persist for the perpetuation of that state.

An example of such a mechanism in prokaryotes is the establishment of the lysogenic state in bacteriophage lambda by cII, a protein that activates transcription of the lambda repressor cI. Once repressor synthesis is initiated, the repressor protein itself binds to a second promoter and activates its own transcription, so that activation from the cII promoter is no longer necessary^{11,12}. An analogous system, in which the function of an enhancer required at one time is later replaced by the function of another enhancer/upstream sequence, seems to operate in the case of the immunoglobulin μ -gene. In conclusion, therefore, previous expression data with weak and strong enhancers⁷, together with the latest data from Grosschedl, Roeder, Green and their colleagues^{2,3,5}, do not support a hit-and-run mechanism for transcriptional activation by enhancers. □

1. Wabl, M.R. & Burrows, P.D. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2452–2455 (1984).
2. Horikoshi, M., Hai, T., Sunlin, Y., Green, M.R. & Roeder, R.G. *Cell* **54**, 1033–1042 (1988).
3. Hai, T., Honikoshi, M., Roeder, R.G. & Green, M.R. *Cell* **54**, 1043–1051 (1988).
4. Klein, S., Gerster, T., Picard, D., Radbruch, A. & Schaffner, W. *Nucleic Acids Res.* **13**, 8901–8912 (1985).
5. Grosschedl, R. & Marx, M. *Cell* **55**, 645–654 (1988).
6. Kelly, D.E. *et al.* *Molec. cell. Biol.* **8**, 930–937 (1988).
7. Schaffner, G., Schirm, S., Müller-Baden, B., Weber, F. & Schaffner, W. *J. molec. Biol.* **201**, 81–90 (1988).
8. Reeder, R.H. *Cell* **38**, 349–351 (1984).
9. Brown, D.D. *Cell* **38**, 359–365 (1984).
10. Mattaj, J.W. *et al.* *Nature* **316**, 163–167 (1985).
11. Wulff, D.L. & Rosenberg, M. in *Lambda II* (eds Hendrix, R.W. *et al.*) 53–73 (Cold Spring Harbor Laboratory, New York, 1983).
12. Ptashne, M. *Genetic Switch* 2nd edn (Cell/Blackwell Scientific, Palo Alto, 1988).

Walter Schaffner is Director of the Institut für Molekularbiologie II, CH-8093 Zürich, Switzerland.

Daedalus

The terrestrial telescope

WE continue to see the Sun two minutes after it has already set: atmospheric refraction bends its rays over the horizon to us. The rays which pass over our heads and escape out into space again are bent a second time as they leave the atmosphere. These rays, which just graze the edge of the spherical Earth tangentially as they go by, are all deviated inwards to form a hollow cone of converging rays. Daedalus calculates that they will meet at a focus 300,000 kilometres on the night side. Seen from this point, the Earth must show a Sun-bright rim of shining atmosphere. It acts as a gigantic lens — indeed, since the rim of the atmosphere focuses about 300,000 gigawatts of solar energy in this way, it acts as a gigantic burning glass.

Rather sadly, no power satellite can be positioned to exploit this torrent of radiation. A satellite can never hover at such a static focus, and would be rapidly vaporized if it could. In a suitable 300,000-kilometre orbit, however, it could sweep through the blazing focus every 19 days. It would have plenty of time to cool off and recover after each traumatic encounter, while transmitting back to Earth the energy it had captured. But Daedalus prefers to see the 300,000-kilometre orbit as the focal plane of the vast refractor telescope formed by the Earth's atmosphere, with the light grasp of a lens 600 kilometres across, and the resolution of one the size of the Earth. A special imaging satellite, its orbit safely tilted away from the deadly solar focus, could scan it.

As the satellite moves, it will see refracted starlight from the other side of the Earth pulsing in the atmospheric rim. The light will be received on a big charge-coupled 'retina' as a bright and hugely magnified image, and accumulated digitally as a scan of the star field swept by the atmosphere/satellite telescope system. Unfortunately the focal length of the atmosphere is longer for its less-refracting outer layers; it suffers from spherical aberration. Its central image may be corrupted by the coma-spread images of off-axis stars. But as any given star pattern moves across the focal plane, first off-axis in one direction, then central, then off-axis the other way, it will exhibit a predictable sequence of such optical distortions. Continuous on-board digital deconvolution of the sequence should be able to reconstruct the undistorted image.

Daedalus's space telescope will scrutinize its swathe of sky as never before. Nearby stars will confess their planets; brown dwarfs will sidle into view; galaxies on the red-shift limit will be resolved into stars. Galaxies beyond it would perturb cosmologists, but please Daedalus. David Jones

Evolution of modern proteins

SIR—In a recent communication Brenner¹ has outlined an approach to understanding the evolution of modern proteins. In particular he has drawn attention to the active serine residue in a number of proteins and suggested that its progenitor was an active cysteine residue (the serine codon AGY being derived from the cysteine codon TGY by a single base change). Brenner also makes the point that the limited catalytic function of primitive peptides might have been enhanced by the binding of metal ions.

It is of interest that the isolation and sequencing of genes for the selenoenzymes mammalian glutathione peroxidase (GSH-Px)^{2,7}, and bacterial formate dehydrogenase (FDH)⁸ has shown that the active-site selenocysteine residue is specified by the 'stop' codon TGA. Also selenocysteine transfer RNA has been shown to be derived by modification of a specific serine-charged tRNA (possibly via a phospho-seryl tRNA intermediate)^{9,10}, suggesting that TGA was originally a sense codon in a primitive genetic code but that this general function was lost on the introduction of oxygen to the biosphere¹⁰.

Thus it is possible that serine was specified by the codon TGN in a more primitive genetic code but that the seryl tRNA could be modified to contain other more catalytically active amino acids such as phosphoserine, cysteine, selenocysteine or tryptophan (TGA is read as tryptophan in mitochondria). Insertion of the appropriate modified serine into the peptide sequence would have been dependent on the surrounding context sequences in the messenger RNA. Thus a diversity of active peptides could have been generated using a restricted amount of genetic information, with further diversification dependent on the evolution of greater exclusivity in tRNA charging, and the development of pathways for the *in vivo* production of complex amino acids.

The present-day retention of the selenocysteine coding function of TGA in GSH-Px and FDH is presumed to be dependent on the surrounding nucleotide sequences in the mRNA. It is of note that in GSH-Px the strongly conserved amino-acid sequence around the active site consists entirely of mono-basic, mono-acidic amino acids, suggesting a primitive origin for this region of the enzyme. In fact the whole GSH-Px polypeptide is deficient in aromatic and heterocyclic amino acids, again suggesting an ancient origin for this two exon polypeptide of relative molecular mass 21,000 (21K). FDH on the other hand, appears to be of more recent origin being larger (70K) and containing a higher proportion of more complex amino acids.

Comparison of the region around the selenocysteine in GSH-Px and FDH does

not indicate any conservation of amino-acid or nucleotide sequence which might in itself be a signal for translation of the TGA codon. However, it is of note that in both cases the adjacent nucleotide sequences form inverted repeats which are themselves preceded by inverted repeats. Whether such structural motifs play a direct role in translation of the TGA codon as selenocysteine, and are indeed of ancient origin, remains to be seen.

Finally, it is of note that in neither GSH-Px nor FDH is the selenocysteine residue associated with the conserved GX SXG or GXCXG motifs¹ found in many enzymes containing active serine or cysteine residues. This would support the hypothesis that in a primitive genetic code TGN coded mainly for modified serine, the particular modified tRNAs being specified by context sequences in the mRNA. Subsequent evolution led peptides containing an active cysteine residue, and its context sequences, down a separate phylogenetic pathway leading, as proposed by Brenner¹, to the active serine enzymes.

PETER S. GOLDFARB

*Molecular Toxicology Group,
Department of Biochemistry,
University of Surrey, Guildford,
Surrey GU2 5XH, UK*

1. Brenner, S. *Nature* **334**, 528–530 (1988).
2. Goldfarb, P. *et al. Nucleic Acids Res.* **4**, 3517–3530 (1983).
3. Chambers, I. *et al. EMBO J.* **5**, 1221–1227 (1986).
4. Mullenbach, G. *et al. Nucleic Acids Res.* **15**, 5484 (1987).
5. Sukenaga, Y. *et al. Nucleic Acids Res.* **15**, 7178 (1987).
6. Ishida, K. *et al. Nucleic Acids Res.* **15**, 10051 (1987).
7. Ho, Y. *et al. Nucleic Acids Res.* **16**, 5027 (1988).
8. Zinoni, F. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 4650 (1986).
9. Sunde, R. & Evenson, J. *J. biol. Chem.* **262**, 933–937 (1987).
10. Leinfelder, W. *et al. Nature* **331**, 723–725 (1988).

Evolution of an active-site codon in serine proteases

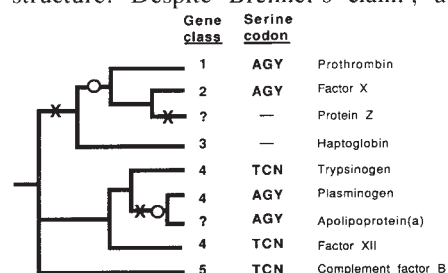
SIR—Brenner¹ proposes that the last common ancestor to vertebrate serine proteases was a cysteine protease that existed billions of years ago, based on the observation that the active-site serine of some serine proteases is encoded by the codon TCN, whereas in others it is AGY. Serine is the only amino acid where it is impossible to change from any one codon to all other possible codons by single nucleotide substitutions without having to pass through an intermediate codon that does not encode serine. I have a simpler explanation for the evolutionary history of this gene family.

The TCN codon for the active-site serine is found in serine protease genes of eubacteria and invertebrates and in some vertebrate genes. In contrast, the AGY codon is found only in serine protease genes of vertebrates, suggesting a late origin from an ancestral serine protease that was of the TCN type. Moreover, the vertebrate AGY proteases are involved in

physiological processes found only in vertebrates²; thus there is no evidence for an ancient lineage associated with these genes.

Molecular phylogenies of the vertebrate serine proteases have been produced³. None of these phylogenies supports Brenner's¹ suggestion that the TCN and AGY proteases represent separate lines of descent from an ancestral protease. Rather, the phylogenies of the genes suggest that the TCN codon evolved on two occasions to an AGY codon, once on the lineage leading to the vitamin K-dependent coagulation factors, and once along the lineage leading to plasminogen and to apolipoprotein(a) (see figure).

Evidence of multiple origins of the AGY serine codon is also found in gene structure. Despite Brenner's claim¹, a



Tree relating selected vertebrate serine proteases and related proteins together with codon used and exon-intron structure. The phylogeny is based on previous alignments³, with the introduction of protein Z (ref. 5) and apolipoprotein(a) (ref. 6) adjacent to the most similar sequences. Gene classes, identified by numbers², indicate genes which share similar intron-exon structure; uncharacterized genes are indicated by question marks. Protein Z and haptoglobin do not have serine at the active site. X denotes the loss of a serine codon and an open circle the gain of a new serine codon. Brenner¹ proposed that haptoglobin and protein Z are descendants of TCN proteases; molecular phylogenies indicate that both are more closely related to AGY proteases.

strict association of gene structure and serine codon is not observed (see figure). Vertebrate serine protease-like genes have been grouped into five classes². The AGY proteases fall into two clusters within the phylogeny of vertebrate serine proteases, with representatives in three classes. The larger group represents the vitamin K-dependent coagulation factors; the second group includes plasminogen and apolipoprotein(a), which are like the trypsin-like proteases. The incompletely characterized gene for plasminogen⁴ seems to have one intron shared with the trypsin-like genes. Thus there seems to be less correlation between active-site codon and gene structure than between amino-acid sequence and gene structure².

My model is based not only on the sequence of the active-site serine codon, but also on amino-acid sequence and gene structure. It seems that the TCN codon has twice been converted into an AGY

codon within the vertebrate lineage, and in one case a descendant of the intermediate non-serine protease (haptoglobin) can be found (see figure). This model also predicts that additional serine proteases may have lost protease function at some point in their evolutionary history, but due to back-mutations (restoring the serine codon to the ancestral type) there is no trace of their non-protease past in their present sequence, in contrast to the AGY serine proteases. The existence of tandem arrays within the family of more than 100 serine protease genes in the vertebrates has probably provided many opportunities for these genes to lose and regain protease function as well as to adapt to new functions.

DAVID M. IRWIN

Department of Biochemistry,
University of California, Berkeley,
California 94720, USA

1. Brenner, S. *Nature* **334**, 528–530 (1988).
2. Irwin, D.M., Robertson, K.A. & MacGillivray, R.T.A. *J. molec. Biol.* **200**, 31–45 (1988).
3. Doolittle, R.F. & Feng, D.F. *Cold Spring Harb. Symp. quant. Biol.* **52**, 869–874 (1987).
4. Sadler, J.E., Malinowski, D.P. & Davie, E.W. in *Progress in Fibrinolysis* Vol. 7 (eds Davidson, J.F., Donati, M.B. & Coccheri, S.) 201–204 (Churchill Livingstone, Edinburgh, 1985).
5. Højrup, P., Jensen, M.S. & Petersen, T.E. *FEBS Lett.* **184**, 333–338 (1985).
6. McLean, J.W. *et al. Nature* **330**, 132–137 (1987).

Universal and essential role of MPF/*cdc2*⁺

SIR—As reported recently in *News and Views*¹, the product of the mitosis-inducing *cdc2*⁺ gene of the fission yeast, *Schizosaccharomyces pombe* is a component of the maturation promoting factor (MPF)^{2–4}, first recognized in *Xenopus* oocytes. This exciting finding demonstrates the universal and essential roles of MPF/*cdc2*⁺ in cell-cycle regulation.

We previously isolated a meiosis-specific recessive mutant *twsl-N22* of *S. pombe* which seems to be normal in meiosis I but unable to enter into meiosis II so that resulting asci contain two viable diploid spores instead of the normal four haploid spores^{5,6}. Mitotic growth of *twsl* is normal at any temperature. More recently we have found that *twsl* is an allele of the *cdc2* gene. Tetrad analysis shows that the *twsl* locus is tightly linked to *cdc2* (less than 1 cM separates them). Cloned *cdc2*⁺ can complement *twsl*, resulting in the formation of four-spored asci. Furthermore, mitotic growth of the *cdc2-33*^{+/twsl} heterozygote is poor at non-permissive temperature, suggesting an interaction between their products or a shortage of the active *cdc2* product in the heterozygote. Temperature-sensitive alleles of *cdc2* also frequently produce two-spored asci at semi-permissive temperature⁷. The *twsl* allele in *cdc2* might be meiosis-specific, such that the mutation specifically impairs a transition step to meiosis II. Alternatively, it might result in a reduced *cdc2*

activity that is still sufficient for mitosis and meiosis I, but not meiosis II. In any case, the phenotype of *twsl* allele has revealed another aspect of the various roles of the *cdc2*⁺ gene in mitosis and meiosis.

OSAMI NIWA

MITSUHIRO YANAGIDA

Department of Biophysics,
Faculty of Science, Kyoto University,
Sakyo-ku, Kyoto 606, Japan

1. Murray, A.W. *Nature* **335**, 207–208 (1988).
2. Dunphy, W.G. *et al. Cell* **54**, 423–431 (1988).
3. Gautier, J. *et al. Cell* **54**, 433–439 (1988).
4. Labbe, J.C. *et al. Nature* **335**, 251–254 (1988).
5. Nakaseko, Y. *et al. J. Bact.* **157**, 334–336 (1984).
6. Nakaseko, Y. *et al. EMBO J.* **5**, 1011–1021 (1986).
7. Hayles, J., Aves, S. & Nurse, P. *EMBO J.* **5**, 3373 (1986).

Licence to slang Copenhagen revoked

SIR—Posiewnik and Pykacz^{1,2} have described an interesting *gedankenexperiment* for which they claimed that the predicted result would depend on whether one employs the Copenhagen or the stochastic (the Einstein–De Broglie) interpretation of quantum mechanics. But even though the experiment seems to be perfectly realizable with optical devices now available (*Physics Letters*), one should not expect a scientific revolution soon.

The idea of the Posiewnik–Pykacz experiment boils down to the following. One performs a single photon Young's double-slit experiment with only one slit open at a time, and tries to get interference. More specifically, one has single photons in a Mach–Zehnder-type interferometer³. Each of the two possible paths of the photon can be blocked by a shutter. The shutters operate so that only one of them can be open at a time (which can be achieved in a relativistically invariant way). Interference can occur only when the parts of the photon wavepacket chopped by the shutters overlap in the detection region. This can be achieved by a suitable lengthening of the interferometer path with the shutter which opens first, which has the effect of delaying the part of the wave representing the photon traversing that path.

This outline description conforms to our intuition, as well as to the Einstein–De Broglie interpretation of quantum mechanics. But despite the claim of Posiewnik and Pykacz, the introduction of the collapse postulate does not change the predictions. The Copenhagen point of view should read as follows. The collapse of a wavepacket represents a gain of information about a quantum object. But the experiment has been devised in such a way that if the photon passes through the shutters, we know only that it has passed through one of them. This is because the two paths are of different optical length. Thus any time correlation between the gate at a shutter and photon's arrival at a detector (placed at the end of both paths)

is wiped out by the time delay. Indeed, interference is a maximum when the time delay is such that from the point of view of the detector, both shutters are 'seen' by it to open and close together.

Posiewnik and Pykacz seem to rely on the simple argument that, because there is a collapse, there can be no interference. But, one must know into what the wavepacket collapses. While, in those versions of Young's experiments presented in Feynman's lectures, the collapse indeed precludes interference, that is not so for the proposition of Posiewnik and Pykacz.

MAREK ZUKOWSKI

Uniwersytet Gdanski, Instytut Fizyki
Teoretycznej i Astrofizyki,
80-952, Gdansk, Poland

1. Posiewnik, A. & Pykacz, J. *Phys. Lett.* **128**, 5 (1988).
2. Maddox, J. *Nature* **332**, 581 (1988).
3. Grangier, P. *et al. Europhys. Lett.* **1**, 173 (1986).

Cystic fibrosis advantage

SIR—In 1982, Quinton¹ presented evidence that anion impermeability is a basic defect in cystic fibrosis (CF), and speculated that "intestinal secretions of persons carrying the mutant CF gene may be less than those of the normal population when stimulated by cholera toxin... Consequently, these persons... should be at some advantage...". This proposal was recently repeated², and several recent findings providing support for Quinton's speculation have also been cited or presented³. But a key finding germane to the argument was not cited. Because CF homozygotes reproduce only rarely, a selective advantage for the CF allele requires that it results in an altered phenotype when present as a single copy in heterozygotes. The laboratories have now demonstrated proportionally reduced fluid secretion by CF heterozygotes for β -adrenergically stimulated sweating^{4,5}. It will be interesting to see if reduced intestinal secretory responses can also be demonstrated for CF heterozygotes, as defective regulation of intestinal secretion in CF homozygotes involves not only cyclic AMP, but also calcium⁶ and cyclic GMP^{7,8}.

JAN BIJMAN

HUGO DE JONGE

Department of Cell Biology and Genetics,
Erasmus University, PO Box 1738,
Rotterdam, The Netherlands

JEFFREY WINE

Cystic Fibrosis Research Laboratory,
Bldg 420, Stanford University,
Stanford, California 94305-2130, USA

1. Quinton, P.M. in *Fluid and Electrolyte Abnormalities in Exocrine Glands in Cystic Fibrosis* (eds Quinton, P., Martinez, R. & Hopfer, U.) 53 (San Francisco Press, 1982).
2. Hansson, G.C. *Nature* **333**, 711 (1988).
3. Baxter, P.S. *Nature* **335**, 211 (1988).
4. Behm, J.K. *et al. Pediatr. Res.* **22**, 271–276 (1986).
5. Sato, K. & Sato, F. *J. lab. clin. Med.* **111**, 511–518 (1988).
6. Berschneider, R.G. *et al. Gastroenterology* **92**, 1315 (1987).
7. de Jonge, H.R. *et al. Pediatr. Pulmonol. Suppl.* **1**, 54 (1987).
8. Bijman, J. *et al. in Exocrine Secretion* (eds Wong, P. & Young, J.) 17–19 (Hong Kong University Press, 1988).

Transatlantic mentalities

Geoffrey Hall

The Mind. Executive editor Richard Hutton. Reporter and executive-in-charge George Page. *Nine-part series on US PBS from 12 October 1988. Co-production of WNET/New York and the BBC.*

The Mind Machine. Series producer Martin Freeth. Presenter Colin Blakemore. *Thirteen-part series on BBC 2 from 13 September 1988. A BBC TV/WNET co-production.*

The Mind. By Richard M. Restak. *Bantam Books, New York:1988. Pp.328. \$29.95.*

The Mind Machine. By Colin Blakemore. *BBC Books, London:1988. Pp.277. £15.95.*

NEXT week, on 7 December, television viewers in New York state will be able to see *The Violent Mind*, a 90-minute 'special' that concludes the nine-part series *The Mind*, presented by television journalist George Page. The series is a co-production involving WNET/13 in New York and the BBC, and a version of this same film, to be screened on 20 December, will also cap the BBC's 13-part series *The Mind Machine*, presented by the Professor of Physiology at Oxford University, Colin Blakemore.

As a selection of real-life dramas, *The Violent Mind* makes for gripping viewing. We are given, among others, the stories of David, a salesman whose motiveless murder of a customer has been attributed to the cerebral effects of the pesticide he sold; of Sandie, whose outbursts of violence coincided with her menstrual cycle; of Julie, who stabbed an innocent bystander in the course of an epileptic episode; and of Jay, a sufferer from manic depression who shot his estranged wife in response to a sequence of hallucinated 'messages'. These stories are told with a meticulous (sometimes almost prurient) attention to detail that commands the viewer's attention. In consequence, one educational function of the film seems certain to be achieved — seeing these case histories is sure to disrupt the complacency of anyone who thinks that to distinguish bad from mad is an easy matter.

When it comes to explanations (that is, to supplying an account of the mechanisms that underlie the behaviour shown by these unfortunate individuals), the film has less to offer. In discussing David's case, for instance, there is some consideration of the possibility that the pesticide he sold might disrupt neurotransmitter systems. But these matters take second place to the question: should David be held responsible for his actions or was his hypothalamus to blame? The psychological and physiological source of Jay's

depressive illness is discussed only briefly; instead we are invited to ponder at length the issue of whether or not the existence of this illness means that his free will was impaired.

It is probably unfair to judge *The Violent Mind* as an orthodox essay in the

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Brainwaves — seen against a background of Blakemore's own EEG and eye movement, are Colin Blakemore on location (top) and Richard Hutton and George Page at the lectern.

popularization of science. The makers of the film explicitly set themselves the task of exploring the relationship between science and the law when the two come together over a violent crime. It was perhaps assumed that the scientific analysis could be carried forward from previous programmes in the series — there has been, after all, a whole film devoted to depression (called *Depression* in the United States, *Portrait in Blue* in the BBC version); and neurotransmitters have been discussed repeatedly, particularly in the programme *Addiction* (*Prisoners of Pleasure* in Britain). It is to the previous films in the series that we

must look for a popular account of the sciences of brain and mind.

It turns out, however, that the earlier films differ from the final one less dramatically than might have been expected. The basic method is still the same — the presentation of intriguing case histories (addicts, psychotics, stroke victims and so on). These cases allow a vivid picture of the issues to emerge, but they are so much natural history and the science starts when we begin an analysis of evidence on and theory about mechanisms. This is much more difficult to popularize and the films have varied greatly in their success (or even in their willingness to make the attempt). That by the BBC on vision (*Sight Unseen*) presented complex scientific issues well; *Pain and Healing* on the other hand, especially in its American version, did little but document the (apparently surprising) fact that the mind

can influence the body.

Perhaps there is some intrinsic limitation to the ability of television to present complex arguments, a limitation that does not apply to the printed word and the still picture. Certainly the authors of the books that accompany the programmes, although these too are intended for a lay audience, are more willing to go into details of experiments and complexities of interpretation. Blakemore himself wrote *The Mind Machine*, but *The Mind* was written by R.M. Restak, a neurologist, who was not involved in writing the television films. However, the basic material in both books closely follows that presented in the corresponding television series.

There are differences between the two books (and the two sets of films, British and American) both in style and in content. The stylistic difference between the books may reflect a difference between the two authors; however that may be, *The Mind* is worthily earnest in its didactic intent, *The Mind Machine* determined to educate by stealth in the course of a good read. But something of the same is seen when we compare British and American television. George Page lectures to us from his studio, letting us know what we are about to see and then summing up by telling us what we have just seen; Colin Blakemore introduces his analogy between the human mind and the Delphic oracle from the foot of Mount Parnassus itself.

However little significance these stylistic differences possess as indicators of national differences in temperament, there is probably even less to be made of

the differences in content. These latter arise largely because *The Mind* was developed as a sequel to a series *The Brain*, shown in the United States in 1984, and thus could not cover topics dealt with earlier. But the BBC was able to include material from *The Brain*, creating new films on schizophrenia, rhythms and sleep, memory and perception. A consequence of this historical accident is that the BBC has been able to put together a more satisfactory series — more satisfactory not only because the coverage of the topic is fuller but also because there is a clearer definition of precisely what the topic is.

The very title makes clear that the British series is devoted to psychological aspects of neuroscience. In the first programme Blakemore defined his subject as "the human mind and the brain that creates it: the mind machine". But the Americans were faced with their own version of the mind-body problem. Having already devoted a series to the brain they found themselves obliged to have a quite separate one on the mind — a dualism that sits uneasily with the assertion made by Page at the start of his series that "the mind is what the brain does". The first programme constituted an

attempt to justify this assertion, and having made the effort, the programme-makers felt entitled to proceed with a discussion of the physiological and cerebral basis of mental development, depression, language and so on.

Not that the makers were fully convinced themselves by the justifications they offered. In *Depression*, for example, although one authority (Whybrow of the University of Pennsylvania) carefully explained that everything that anyone does has a neurochemical equivalent, we were still presented with a discussion of the difficulties involved in distinguishing psychologically caused depression from depression that is biological in origin. The viewer might be forgiven for taking away the impression that mental states are sometimes 'what the brain does' and sometimes they are something else.

Perhaps these problems would have been avoided if *The Mind* had been faithful to its title, and had concentrated on the psychological analysis of perception, mental illness, language, thinking and the rest. Experimental investigation of these topics has generated a set of well-formulated psychological theories (a psychological theory being a fairly abstract specification of the mechanisms that must exist in order to generate the behaviour that is observed). Concentrating on this approach does not mean that consideration of cerebral mechanisms must be omitted. Psychological and physiological approaches interact. Psychological theories have sometimes been described as constituting descriptions of a 'conceptual nervous system'. Some of the features of the conceptual system can be identified in the real nervous system. Others, as yet, cannot; but the process of trying to make the identification frequently proves fruitful in developing our understanding of both real and conceptual systems.

Finally, the scientific community in Britain should offer a vote of thanks to Colin Blakemore. For some time he has been a target for the attacks of extreme campaigners against 'vivisection'. By presenting *The Mind Machine*, which includes examples of experimental work on non-human animals, he is likely to attract yet more unwelcome attention — indeed, on 8 September *The Sun* newspaper printed a hostile piece even before the BBC series began ("This telly rat must be stopped"). It is only by carefully explaining the purposes of experiments on animals to a wider audience that public sympathy can be won. The hazards that follow public exposure (abuse from a minority) fall to Blakemore alone; the advantages in terms of a heightened understanding by the majority accrue to us all. □

Geoffrey Hall is in the Department of Psychology, University of York, York YO1 5DD, UK.

What's doing the rounds?

T.H. van Andel

Chemical Cycles in the Evolution of the Earth. Edited by C. Brvan Gregor, Robert M. Garrels, Fred T. Mackenzie and J. Barry Maynard. Wiley: 1988. Pp.276. \$44.95, £39.95.

THIS book purports to present a coherent overview of the subject suggested by the title. In reality, it consists of a set of articles which, as seems inevitable in volumes of this kind written by a large number of authors, are loosely and not very systematically related to it. The main contributions are preceded by a "bibliographic note", a memorial to the late Dutch mineralogist Willem Nieuwenkamp to whom the volume is dedicated.

A prologue, announced as a sketch of the development of thought over the past two centuries on cyclic processes in geology, turns out to be an oddly assorted collection of historical notes. Because there is little evidence for such thought until very recently, the author has been forced to adopt a quite peculiar perspective. The approach has little historical merit, and except for the many scattered eyebrow-raisers we are told little that is new or even relevant.

The body of the book consists of chapters on the geochemical cycles of the continental crust and ocean system, and of atmospheric gases, on the interaction between external geochemical cycles and the mantle, and on the biogeochemical cycles of carbon and sulphur. Although the word "cycle" is used on almost every page, the concept of geological cyclicity is nowhere rigorously defined, and its usage in the various contributions is so wide ranging that it obscures rather than helps, and often merely appears to be a nod in the direction of current fashion.

In virtually every chapter the style and level of treatment oscillate between a rather self-conscious rigour, with a good deal of unnecessary mathematics, computation and chemical equations, often sliding over into the densely obscure, and discourses of first principles that would be more appropriate in an undergraduate textbook. The referencing illustrates the point: the sources of key principles and basic facts (such as the composition of the atmosphere) are often taken to be college textbooks, and on p. 183, for example, a popular National Geographic book is cited as the reference for the ages of the ocean floor as derived from magnetic anomalies. Also, apart from the usual excessive citations of the authors' own works, the bibliography is oddly selected and several important currents in exogenic

ADVERTISEMENT

Up To 100 Tissues On One Slide!

Get 20 to 100 normal or tumor tissues on **ONE** slide for monoclonal antibody screening.

Combinations of tissues, 5 kidney, 5 breast, 5 lymph nodes etc. etc., are grouped on each slide so that a spectrum of tumor or normal tissues are tested at one time.*

These paraffin fixed tissues speed monoclonal testing, and save immeasurable time and money.

For information about **Multi-Tissue Technology** and the slides that are immediately available, contact:

Xenetics Biomedical, Inc.
4330 Barranca Pkwy./Suite 101-364
Irvine, California 92714
(800) 888-4999 or (714) 651-1245

*For research purposes only, not for diagnostic use.

geochemistry are more or less ignored. Overall, the book carries a strong flavour of the 'school' of Garrels and Mackenzie, innovative and interesting at one time, but that time was nearly 20 years ago.

Consequently, several evenings of rather laborious reading failed to provide this reviewer with much additional insight into the subject. I make an exception for the final chapter, as yet unmentioned, an interesting and usefully controversial comment by Jan Veizer on the evidence for cyclicity as he sees it manifested in detrital sediments. The treatment here is lively and accessible, although somewhat vitiated by an unfortunate and rather unhelpful insistence on a division between evolutionist and cyclist perspectives on the matter.

Still, no coherent overview emerges of the magnitude of geochemical cyclicity in the history of the Earth, nor does the book necessarily persuade the sceptic that the concept is worth taking seriously. It is also far from clear what the intended audience

is, as is so often the case with the multi-authored 'syntheses' that are springing up across the literature like toadstools in the autumn. The contributions are either too short and obscure to be introductions to the field for outsiders, or are too elementary to spark an interest in the specialist.

The test of a volume such as this is whether its whole is better than the sum of its parts; I would argue that in this case we would have been better served if the few segments of original writings had been presented in suitable journals (which, incidentally, also would have ensured better editorial quality control). As it is, the book will rest on many a library shelf after having consumed monies that might have been used better for other works. I doubt that it will be widely consulted, or will be found to have been a source of inspiration after another decade of research has passed. □

T.H. van Andel is in the Department of Earth Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EQ, UK.

Clash of cultures

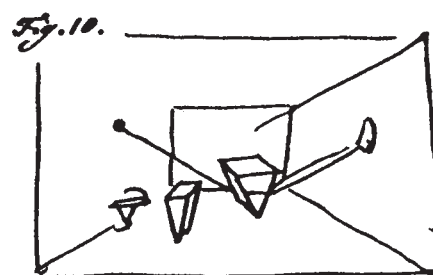
J.D. Mollon

Goethe contra Newton: Polemics and the Project for a New Science of Color. By Dennis L. Sepper. Cambridge University Press: 1988. Pp. 222. £27.50, \$39.50.

How astonished I was, then, when the white wall, observed though the prism, remained white just as before; that only there, where darkness adjoined on it, did a more or less distinct color appear. . . . It did not take much reflection for me to recognize that a boundary is necessary to produce colours, and I immediately said to myself, as if by instinct, that the Newtonian teaching is false.

It is this unashamed admission, placed at the end of the *Farbenlehre*, that has always coloured the response of the scientific establishment to what Goethe himself saw as his greatest contribution. Under pressure to return the borrowed prism, Goethe made the most casual and hurried observations, wrongly supposed that newtonian optics were contradicted by what he saw, and set out on a vast, vain attempt to build a rival synthesis of colour science. For the poet and *Literat*, white could not be a composite: nature is not to be rent by torturing experiment and abstraction, just as a work of art should not be subjected to an analysis that destroys it.

Sepper's purpose is to show us that this caricature is inadequate. Goethe, he argues, had a subtle and well-considered view of the nature of scientific method, had good grounds for questioning Newton's implicit presuppositions, and had even better grounds for criticizing the



A drawing from Goethe's notebooks showing what he called the "objective" and "subjective" ways of using the prism: a newtonian spectrum is being formed objectively and is here being viewed subjectively through another prism.

degenerate accounts of newtonian optics that are found in eighteenth-century compendia of physics. To illustrate Goethe's method, Sepper gives a detailed summary of the first of the *Beiträge zur Optik* (1791), which describes an ordered sequence of experiments on the appearance of edges when viewed through prisms. For Goethe, the purpose of experiments is to allow the phenomena to show themselves, in a way that is controlled but is uncontaminated by hypothesis or by the observer's more general *Vorstellungsart*. A single experiment cannot stand on its own: the conditions must be systematically varied — contrasted, simplified, recomplified — to show the relationship of one phenomenon to another. This is the manifolding of experiments that Goethe calls *Vermannigfaltigung*. Sepper draws our attention to the essay *Der Versuch als Vermittler von Objekt und Subjekt* (1793), which makes explicit the scientific method mostly only implicit in the earlier *Beiträge*.

One large section of Sepper's book does seem out of place, and that is the part given over to knocking Newton. The points that Sepper makes against Newton

(for example, the ambiguous status of the ray — as an observed phenomenon and as a theoretical construct; and the inadequacy of the *experimentum crucis* as a refutation of Hooke) are indeed in the spirit of Goethe, but in their detail they owe more to the analyses published by newtonian scholars since 1960.

Sepper explicitly sets up Newton and Goethe as contemporary rivals, and this is artificial. Colour science did not stand still during the eighteenth century, and many concepts and observations were available to Goethe at the publication of the *Farbenlehre* in 1810 that were not available to Newton in 1672. Goethe is given much credit by Sepper for his survey of 'physiological' colours (for example successive and simultaneous contrast); but these phenomena are richly described in the eighteenth-century literature, and by 1810 the better commentators had firmly made the categorical distinction between those aspects of colour that needed physical explanation and those that needed physiological explanation. (It was a newtonian physicist, Lichtenberg, who first clarified the distinction for Goethe — see H. Lang *Photoin* 6, 12–31; 1983.)

Against Newton and in favour of Goethe, Sepper cites (pp. 14 and 156) the modern demonstrations of colour constancy by Edwin Land, which show us (or rather, remind us) that there is no fixed relationship between the refrangibility of light and the hue perceived. But as I have related in *The Listener* (113, No. 2891, 6–7; 1985) the latter observation had already been made, and the nature of colour constancy lucidly explained, by the establishment scientist, Gaspard Monge, two decades before the publication of the *Farbenlehre*. Monge grasped the relationship between colour constancy and coloured shadows, and he proposed, as Land did a long time later, that hue depends on the ratio of different components in the spectral flux, rather than upon their absolute values. The paper that Monge gave to the Royal Academy of Sciences in 1789 was well known to his contemporaries; and we know, from a reference in Goethe's notebooks, that Goethe was aware of it (*Goethes Werke, Abtheilung II: Naturwissenschaftliche Schriften Band V: Paralimpomena*, p.132; Weimer, 1906).

. . . and one does not stop to wonder, through what incredible lapse of thought so outstanding a man not only deludes himself at the outset, but lets the error so take root, that against all evidence, indeed against all conscience, he clings to it thereafter and invents one unseemly experiment after another, in order to hide his initial inattention from his inattentive pupils [Goethe, *Farbenlehre, historischer Teil, achtzehntes Jahrhundert: "Isaak Newton"*].

J.D. Mollon is in the Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK.

Booking the cooks

Kenneth Mellanby

But the Crackling is Superb: An Anthology on Food and Drink by Fellows and Foreign Members of the Royal Society. Edited by Nicholas and Giana Kurti. *Adam Hilger: 1988. Pp.261. £12.50. Available in the United States from Taylor & Francis Inc., \$29.*

PROFESSOR Nicholas Kurti, an authority on low-temperature physics, deserves our condolences, in that for his new anthology of kitchen hints he has failed to goad his distinguished fellow contributors to reach his own high standard of amiable culinary eccentricity.

In recent years, the fashion has grown up of what one might call 'instant books'. You wish to compile a book on dogs so you write to a large number of well-known people asking about their pets. A few reply and hey presto! you have a best-seller. This is a variant of the technique of standing in a shopping mall with pencil and notebook asking people impertinent questions about their sex lives, their religion or their choice of the best household detergent. Alas, these approaches seldom produce a genuine book of value or coherence — *But the Crackling is Superb* is no exception.

What is rather perplexing is that Professor Kurti's gathering ground was so fertile. He wrote to several hundred fellows of the Royal Society — the cream of Britain's scientific intelligentsia — asking for recipes or food advice, but one really has to admit the results do not bolster anyone's reputation. If a distinguished physicist can produce a recipe of such monumental imprecision as at least one of the pieces in this book, one finds oneself wondering what his scientific experiments can be like. No editor of *Women's Institute* anthology, or of a village collection of local recipes, could pass by contributions that are so inexact, or commonplace, or both.

Unsurprisingly, Professor Kurti's own contributions combine expertise with innovation. There certainly is a place for science in the kitchen, and he leaves us with a delightful picture of himself sticking several thermometers into a *soufflé*. I personally admire his creative approach and share his enthusiasm. In years gone by my interests as given in *Who's Who* included wine and food, until advancing years, a lowered capacity for intake and, above all, growing awareness of ecological constraints, obliged me to change the entry into "austere living".

So instead of passing learned comments on food and drink I now shall have recourse to yet another technique for producing instant books — that is, anthologizing from anthologies. I rather like a

quotation from Lord Horder, once medical adviser to Britain's royal family:

Our bodies are not puritanical. The pleasant habits of eating and drinking were never meant to be subject to a chemical equation.

Perhaps Professor Kurti would disagree, and prefer the surgeon Sir Robert Hutchinson:

Vegetarianism is harmless enough, though it is apt to fill a man with wind and self-righteousness.

Kurti is certainly not a vegetarian. Let us try Magnus Pyke, probably Britain's best-

known nutritionist:

Meat is a status dish in which the sizzle counts for more than the intrinsic nutritional worth.

But as a staunch ecologist I would finally echo another piece of wisdom from Magnus Pyke:

We are wealthy and wasteful but this can't go on. If we don't eat dog biscuits, we could end up eating our dogs instead.

Kenneth Mellanby, 38 Warkworth Street, Cambridge CB1 1EG, UK, is a keen cook and is author of Can Britain Feed Itself? (1975).

New angles on momentum

R. N. Dixon

Angular Momentum: Understanding Spatial Aspects in Chemistry and Physics. By Richard N. Zare. *Wiley:1988. Pp.349. \$39.95, £34.50.*

WE ARE all familiar with phenomena that are related to angular momentum: that a bicycle at speed does not readily fall over, and that the currents in the oceans of the rotating Earth tend to follow a circular motion. We also know that the electrons in an atom tend to revolve around the nucleus as in a miniature Solar System. Thus all chemistry and physics undergraduates learn of the importance of angular momentum quantum numbers in describing the motions of atoms and molecules. Even so, the mathematical description of angular momentum baffles most of us, not least because the formal examples usually found in textbooks seem so remote from practical application. Richard Zare's book is destined to change all that.

A fundamental aspect of the layout of the book is that the basic theory introduced in each chapter is put to use when the student solves the problems at the end of the chapter. These are no mere formal exercises: they illustrate how the theory works in explaining such widely different experiments as the anisotropy of the scattering of beams of atoms fired at one another, and the nuclear spin structure in the microwave spectrum of the C₁₀N molecule. In the words of the author, "for someone to read the text and skip the problem sets and applications is like someone reading a book on how to play the piano without ever sitting before a keyboard".

The book is intended for an American one-semester graduate course. It is equally appropriate for a graduate course in Britain, particularly for atomic physicists, physical chemists or chemists using spin-dependent analytical techniques such

as magnetic resonance. It may even find a place in specialist courses in the final undergraduate year. It requires the student to have only a basic knowledge of quantum mechanics, and has the merit of leaving in many of the intermediate steps in the derivations.

Zare starts with the angular momentum of a single quantum particle, then generalizes the theory to two or more coupled angular momenta. Transformations under rotation are described in detail, and applied to separating the physical and dynamical parts of molecular problems from the angular or geometric parts. Applications include the polarization of fluorescence from rotating molecules, and quantum beats. This leads on to the energy-level structure and wave-functions of a rigid rotor. An appendix contains a FORTRAN listing of a program for calculating numerical values of all the angular coupling factors which appear in practical problems. The four problem sets and 16 applications make up one third of the length of the volume.

This is certainly not the first book on the theory of angular momentum, but it is the first to combine the basic theory with excellent illustrations of how that theory works across a wide range of topics in chemistry and physics. It will be widely used as a learning text for postgraduate students, and as a revision primer and reference work for their tutors. □

R. N. Dixon is in the School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK.

New editions

- *Soil Physics*, 2nd edn, by T. J. Marshall & J. W. Holmes. Publisher is Cambridge University Press, price is hbk £45, \$89.50, pbk £15, \$29.95. For review see *Nature* **289**, 731 (1981).
- *Supercomputers and their Use*, revised edn, by C. Lazou. Publisher is Oxford University Press, price is hbk £25, pbk £12.50. For review see *Nature* **326**, 910 (1987).
- *Pioneers of Science: Nobel Prize Winners in Physics*, 2nd edn, by R. L. Weber. Publisher is Adam Hilger, price is £12.95.
- *Ecotoxicology: The Study of Pollutants in Ecosystems*, 2nd edn, by F. Moriarty. Publisher is Academic, price is Hbk £22.50, \$44.50, pbk £10.95, \$23.

Deleterious mutations and the evolution of sexual reproduction

Alexey S. Kondrashov

Research Computer Center, Pushchino, Moscow Region 142292, USSR

The origin and maintenance of sexual reproduction continues to be an important problem in evolutionary biology. If the deleterious mutation rate per genome per generation is greater than 1, then the greater efficiency of selection against these mutations in sexual populations may be responsible for the evolution of sex and related phenomena. In modern human populations detrimental mutations with small individual effects are probably accumulating faster than they are being eliminated by selection.

SEXUAL reproduction—the alternation of meiosis and syngamy with attendant segregation and recombination—is one of nature's wonders (Fig. 1). Individuals who generally compete for existence cooperate in the key process of reproduction. With obligate asexual reproduction, as in many prokaryotes¹ and in some eukaryotes, such as the dandelion *Taraxacum officinale*, which probably lost sex only recently, every individual has only one parent and all the genes in the genome are permanently linked together. Sexual reproduction which is predominant among eukaryotes, disrupts this linkage even when sporadic, and the gene *sensu* Johanssen (the unit of function) becomes equivalent to the gene *sensu* Mendel (the unit of heredity). It also leads to important characteristics of individuals and populations, such as frequency of genetic recombination and mode of mate choice, which are absent in asexual populations.

Although the investigation of genetic segregation and recombination, both direct results of sex, was initiated by Mendel, the evolution of these processes was neglected until the 1930s, probably because biologists were satisfied with a general explanation proposed by Weismann more than a century ago² (see also ref. 3). By the early 1970s it was appreciated, however, that asexual reproduction has an intrinsic twofold advantage over anisogamous sexual reproduction (the most common in higher organisms). In an asexual population of stable size, each individual produces an average of one progeny, whereas in a sexual population with a 1:1 sex ratio each female produces an average of one male and one female progeny. Hence, if a mutation appears causing females to produce two asexual female offspring, its frequency will double in each generation³⁻⁹. Furthermore, the increased complexity associated with the sexual mode of reproduction imposes additional inherent costs^{9,10}. So what maintains sex, despite the large advantage of asexual reproduction? This has led to an explosive growth in the number of theories trying to resolve the 'problem' of sex^{5,6,10,11}.

It has been tempting to look for non-evolutionary explanations: for example, sexual reproduction could have some physiological advantage; or sex, having originated for some purpose in the past, exists now only as a vestige; or sexual reproduction, although inherently disadvantageous, is maintained as a by-product of another useful process, such as meiotic DNA repair¹² or biased gene conversion¹³.

Several observations contradict these explanations. As a physiological means of self-propagation, asexual reproduction is quite efficient in many species from various taxa^{5,6,10}. Sometimes individuals from the same population can reproduce both sexually and asexually (facultative sex), which could readily lead to the rapid fixation of obligate asexual reproduction. Sex must therefore be maintained by some strong factor that continues to operate. This factor must not only provide an advantage for the whole sexual population, but also operate in terms of

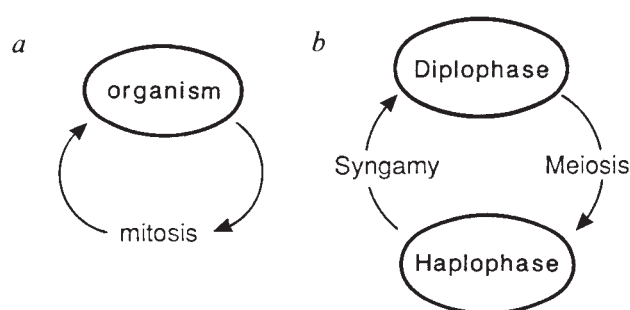


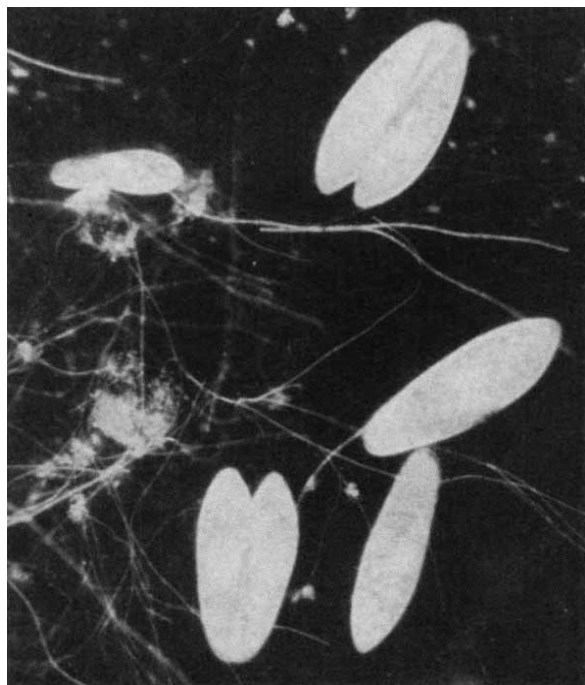
Fig. 1 Life cycle with *a*, asexual and *b*, sexual reproduction. With asexual reproduction each new organism originates from mitotic cell division, which does not change the genotype. With sexual reproduction the life cycle consists of two parts. Diplophase begins with syngamy, that is, fusion of two haploid gametes, so that each cell in this phase contains two sets of genes. Haplophase begins with meiotic cell division, which involves a halving of the amount of DNA as well as genetic recombination from independent segregation of nonhomologous chromosomes and crossing over between homologous chromosomes. In some taxa, for example some plants and fungi, multicellular organisms can be produced by mitosis from either phase to give either haploid or diploid individuals. In other taxa, however, syngamy immediately follows meiosis, as in animals, or meiosis immediately follows syngamy, as in some algae, so only one phase is represented by a multicellular organism.

individual intrapopulation selection⁵. Some processes for which sex or recombination have been claimed to be an unavoidable concomitant can function well without them, as exemplified by meiosis in *Drosophila melanogaster* males, which is not normally accompanied by crossing over^{5,10,11}.

With many other authors^{5,6,10,11}, I argue that sexual reproduction must enjoy some evolutionary advantage. This means that the advantage is not caused by the process itself but by the changes it causes in progeny genotypes (as a result of recombination), which should drive the evolution of sex.

Survey of hypotheses

In 1887 Weismann proposed that sex is advantageous because it is 'a source of individual variability furnishing material for the operation of natural selection'². Some data suggest that sexual reproduction can actually cause enhanced fitness of at least a portion of the progeny^{14,15} but the mechanism of this is obscure. Any evolutionary explanation for the maintenance of sexual reproduction can probably fit into Weismann's framework, because sex does not immediately change allele frequencies and consequently cannot directly improve the population.



The ciliate protozoan *Paramecium* can reproduce both asexually and sexually. Here two pairs are conjugating, the sexual process during which there is an exchange of genetic material between individuals. Photo: John Walsh.

The question remains as to precisely what kind of selection is facilitated by sex. There are two ways of classifying the hypotheses proposed for the evolution of sex: first, what changes (environmental or genomic) lead to selection which sex is supposed to facilitate, and second, by what mechanism (deterministic or stochastic) are the genetic consequences of sex supposed to facilitate selection. In all populations the prevalent genotypes match the environmental demands to some degree. The better the match, the better the adaptation. Sex, if it is to be advantageous, must maintain a better genotype-environment match than that under asexual reproduction. If this match is perfect in an asexual population, sex would be harmful, because it destroys perfect gene combinations and gives rise to segregation and recombination loads^{5,10}. The match can be impaired by either environmental changes or errors in genome copying. Hence, two types of hypotheses, environmental (ecological) and mutational (genetical), are possible.

Introduction of sexual reproduction will not necessarily improve this imperfect match in an asexual population. What sex does is to reduce correlations between the population distributions of alleles at different loci—in other words it destroys linkage disequilibria. Naturally, this randomization of population genetic structure will be advantageous only when it increases the frequency of genotypes with many useful alleles. If some alleles are advantageous in one genetic background and deleterious in another, recombination will be harmful because it destroys coadapted gene combinations^{5,10}. In other words, under conditions where sex would have an advantage, in an asexual population favourable alleles should be distributed more uniformly than at random, leading to lack of both 'very good' and 'very bad' genotypes compared with linkage equilibrium. In this case sex may be favourable both at population and individual levels, as 'very good' genotypes that are sexual in origin reproduce more readily, leading to the spread of sex.

Existing hypotheses attribute the suggested lack of favourable genotypes in asexual populations either to stochastic events in

finite populations or to deterministic factors that would also operate in very large populations^{10,11}. So, four types of hypotheses about the evolution of sex have been proposed: stochastic and deterministic for both environmental and mutational.

My purpose is to consider mutational hypotheses, so I will not review the numerous and varied environmental hypotheses in detail. The advantage of sex has been attributed to increasing differences between either sibs or parents and offspring, which are advantageous under some circumstances, or to increasing the ability of the population to track various kinds of environmental fluctuations. The occurrence of favourable mutations implies environmental changes, so hypotheses that propose that the advantage of sex is a result of increasing the efficiency of directional selection, starting from single advantageous alleles^{5,16} or their low initial frequencies¹⁷, should be classified as environmental (see refs 10 and 11 and the other chapters from the same book for reviews).

Some environmental hypotheses do not contradict our knowledge of population biology, even though conditions providing the more-than-twofold advantage of sex necessary to offset its twofold disadvantage are restrictive^{18,19}. But each environmental hypothesis relies on specific assumptions about environment and can hardly explain sex in all the species in which it occurs. The assumption that all sexual populations undergo fast evolutionary changes or live in environments sharply fluctuating in space and/or in time is unreasonable. Many living fossils which have changed very little in hundreds of millions of years, as well as the inhabitants of very stable and uniform deep-sea ecosystems, are obligate sexual reproducers. Some experimental data also do not support these hypotheses^{20,21}.

Mutation hypotheses are free from this drawback because mutability is an inherent feature of the gene and most non-neutral mutations are deleterious. Two such hypotheses have been proposed. The first of these is a stochastic hypothesis known as Muller's ratchet²² (see also refs 5, 10 and 23–25). In a finite asexual population under the pressure of deleterious mutations, Muller noted that a random loss of all mutation-free individuals is irreversible, whereas with sexual reproduction those genotypes would be re-established by recombination. This mechanism, however, provides a large advantage of sex only in small populations^{5,25}, because with a large population the random loss of the best genotype becomes improbable. Moreover, this hypothesis can explain only the disadvantage of obligate asexual reproduction, but not the evolution into a sexual population.

Until recently, in studies of the evolution of sex and recombination it has always been presumed that, in a very large or infinite population, selection against mutations works independently of the reproduction mode^{5,23–27}. Because this is not really the case, the deterministic mutation hypothesis, with which I will now be concerned, was developed.

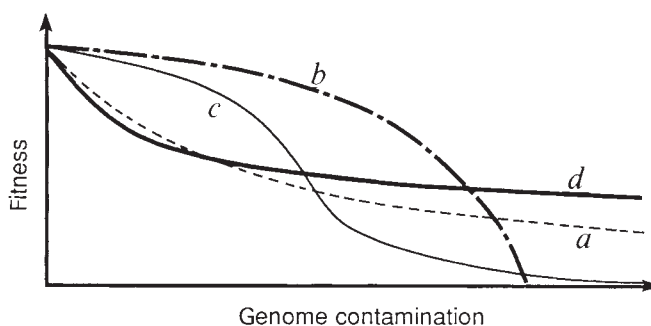


Fig. 2 Various types of selection against mutations: *a*, exponential selection (no fitness interactions between mutations); *b* and *c*, positive interaction, that is, each additional mutation leads to a larger decrease in relative fitness; *d*, negative interaction.

Deterministic mutation hypothesis

Kimura and Maruyama²⁸ have shown that in a sexual population the mutation load depends on the mode of selection and may be much smaller than in an asexual population, where the proportion of individuals that do not reproduce as a result of selection against mutations (L , mutation load^{10,29}) is always $(1 - e^{-u})$ under mutation-selection equilibrium, where u is the deleterious mutation rate per genome per generation. They did not apply this idea to the evolution of sex, however, probably because at that time mutation rates were thought to be small, even per genome (see below).

The condition for $L_{\text{sex}} < L_{\text{asex}}$ is a positive epistasis²⁸⁻³³ under which each additional deleterious mutation leads to a larger decrease of relative fitness; in the extreme case it gives rise to truncation selection under which individuals carrying more than some specific number of mutations do not reproduce at all. Under truncation selection or something similar, L_{sex} could be small even with large u , and the advantage of sex $(1 - L_{\text{sex}})/(1 - L_{\text{asex}})$ can be more than twofold if $u > 1$, as $L_{\text{asex}} > 0.5$ with such u . With asexual reproduction deleterious mutations are eliminated separately according to Muller's principle, 'one mutation, one genetic death'³⁴. On the other hand, sex with truncation or similar selection obviates this principle because the genotypes that are eliminated can contain many mutations³¹, which may give a sexual population an enormous advantage. This is the deterministic mutation hypothesis of the evolution of sex^{32,35,36}.

The underlying mechanism is very simple²⁸⁻³³. We will designate the number of deleterious mutations in the genome as genome contamination, assuming for simplicity that all mutations confer the same deleterious effect. Mode of selection is a function relating fitness of the individual to its genome contamination (Fig. 2). Selection with positive epistasis decreases the variance of the population distribution of genome contamination, p . In a sexual population recombination restores the variance and maintains a roughly independent distribution of alleles at different loci. This means that p is close to normal (as the limit of the Poisson), with the variance σ^2 equal to the mean c . On the other hand, in an equilibrium asexual population p has a much smaller variance; in other words, an asexual population is deficient in very good and very bad genotypes. So, sex enhances the efficiency of truncation selection by maintaining a larger variance (Fig. 3).

At equilibrium the decrease of c resulting from selection is equal to its increase resulting from mutation, u , and the mutation load depends mainly on $v = u/\sigma$. I have called this quantity the genome degradation rate³³ as it takes into account both mutation rate and the deleterious effect of the mutants. Under truncation-like selection, even with large c , the fitness of many individuals may be about as high as that of mutation-free ones. In this case, with small v the mutation load is close to zero. As v increases, the mutation load also increases, approaching 1 when $v > 2$. This is because only a small proportion of individuals have contaminations deviating by more than 2σ from the average, and selection decreasing c by 2σ must eliminate a large majority of the population (Fig. 3).

The situation seems paradoxical: when c , and consequently σ , increases (because $\sigma \approx \sqrt{c}$), the mutation load necessary to counterbalance a given mutation rate decreases. With sufficiently large c , a sexual population can be at equilibrium for any given u with an arbitrarily small L . Even with large u , some progenies in sexual populations (in contrast to the asexual ones) receive less contaminated genomes than the genome of either parent, and under a given u the proportion of such progeny increases with the increase of c , which leads to more effective selection, provided the shape of the selection curve does not depend on c .

This comparison of sexual versus obligate asexual reproduction is relatively simple, as we have to consider only the overall population characteristics^{32,36-38}. When some form of sexual reproduction is established, however, evolution of features of

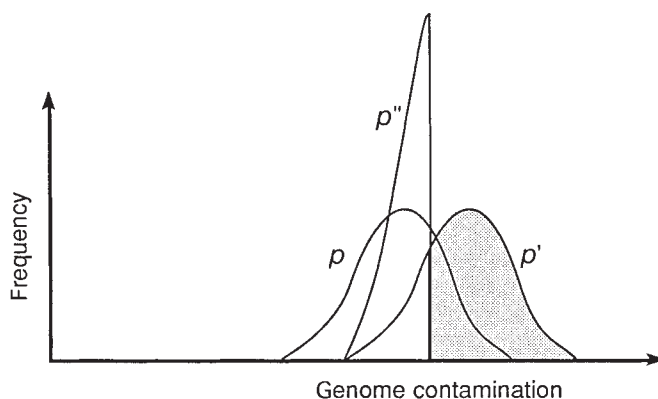


Fig. 3 Scheme of the factors acting in a sexual population at equilibrium under mutation pressure. During one generation mutation, selection and reproduction all occur. Distribution of genome contamination before mutation is p . Mutation shifts it rightward by u (p'). As a result of selection, individuals corresponding to the shaded part of p' die and the distribution after selection is p'' . Reproduction restores normality and the rough equality of mean and variance but does not change the mean. At equilibrium the resulting distribution must coincide with p , so that the mean values of p and p'' are equal. The decrease of average contamination because of selection is proportional to the standard deviation of p' , σ , and, of course, depends on the mutation load: elimination of, say, 50% of individuals with the most contaminated genomes leads to decrease of average contamination by 0.8σ .

sexual reproduction such as recombination frequency, mate choice and inbreeding avoidance becomes possible. The large heritable variance of crossing-over³⁹ and inbreeding⁴⁰ rates show that these characters could readily be selected. To address such evolution, as well as the problem of the origin of sex theoretically, one needs to study intrapopulational processes, considering selection at a modifier locus influencing some feature of reproduction.

These problems have given rise to much controversy. Reduction of recombination frequency leads to decrease of recombination load^{3,5,10,11}; mate choice is of no advantage when the population reaches the selection optimum, as fitness heritability is zero in this case⁴¹, and close inbreeding may be advantageous as it reduces the cost of meiosis^{9,42}. Besides, crossing over may decrease the fitness of an individual in which it occurs⁴³. Hence, although it has been possible to demonstrate potential advantages of these processes^{11,17,42,44}, the problem remains. Nevertheless, crossing over exists in all sexual species, although sometimes in only one sex; mate choice is also quite frequent, and there are various mechanisms for avoiding inbreeding^{6,42,45,46}. In some cases an increase in progeny fitness resulting from mate choice has been demonstrated^{45,46}. Theoretical investigations show that the deterministic mutation hypothesis can be applied to these problems. Crossing over^{47,33}, mate choice⁴⁴⁻⁴⁶, inbreeding avoidance (outcrossing)⁹ and genetic transformation⁴⁸ may be advantageous if the population is at the mutation-selection equilibrium with positive epistasis.

Evolution of obligate versus facultative sexual reproduction needs a similar approach. In this case selection favouring obligate sex is much weaker and appears under more restrictive conditions than in the comparison of obligate sexual versus asexual reproduction: the gene pools of the sexual and asexual subpopulations are not completely separated, so the advantage of sex cannot accumulate in successive generations. The deterministic mutation hypothesis can, however, explain the existence of various forms of facultative sexual reproduction, cyclical included^{9,49}.

In the evolution of most of these phenomena, v is decisive. With $v < 0.5$, selection against mutation does not influence the

evolution of features of reproduction within a sexual population^{9,33,45}. With $v > 0.5$, a modifier allele increasing mate choice⁴⁵ or crossing-over rate³³ enjoys a considerable advantage. I have probably underestimated this advantage for crossing over in the multichromosome genome³³. Note that selection does not favour the increase of the crossing-over frequency beyond 1–2 per chromosome, which is in good accord with the fact that the chromosomes of most species studied have genetic maps of relatively constant length, no longer than 200–300 centimorgans⁶, although the amount of the DNA per chromosome varies by several orders of magnitude.

Avoiding inbreeding completely may be advantageous only when deleterious mutations are recessive to some degree. The whole picture may be rather complex: under the same conditions, for example, both obligate outcrossing and a high rate of self-fertilization may be stable⁹. The increase of the outcrossing rate²⁴ and the chiasmata number per chromosome⁵⁰ in species that live longer can easily be attributed to the possible increase of u and v in these species.

Obligate sex becomes established instead of facultative sex if $v > 1.25$, when the twofold cost of sexual reproduction is considered⁹. A similar approach could be applied to the evolution of gender, or mate types, as mating between two gametes produced by the same haploid individual is genetically equivalent to apomixis. Probably the more stringent conditions necessary for the evolution of obligate sexual reproduction and obligate outcrossing, compared with those leading to the evolution of crossing over, are responsible for the observation that facultative asexual reproduction and inbreeding are very common, but there are no known sexual species without crossing over. Because with $v > 2$ the mutation load becomes intolerable³³, we can conclude that $0.5 < v < 2.0$ in all sexual populations provided, of course, that selection against mutations really is a leading factor in the evolution of reproduction, both at the inter- and intrapopulation level. The next three sections consider the evolution of mutability and data on u and v , respectively, which are of primary importance to the hypothesis.

The evolution of mutability

If we assume that most non-neutral mutations are deleterious, selection will favour the unlimited decrease of mutation rate. Quantitative considerations show that in asexual populations this selection is stronger than in sexual ones, because in the absence of recombination a modifier allele that increases mutation rate will always stay with a contaminated background. The strength of selection for reduction of mutation rate in a sexual population depends mainly on v , being small when $v < 0.5$ and growing rapidly as v increases³³.

What prevents zero mutability being established? If we do not consider rare favourable mutations, the only factor remaining is the physiological cost of high fidelity in the handling of DNA, which would lead to decrease in the fitness of individuals producing less contaminated progeny⁵¹. If such a cost really exists and is high enough, a large mutation rate is to be expected. The two main sources of mutations are mistakes in DNA reparation and in replication. It has become evident that the maintenance of the integrity of DNA molecules in the cell is very complicated^{52–54}. Long DNA molecules are very unstable even in the absence of such 'unnatural' agents as radiation and chemical mutagens. According to Vilenchik⁵⁵, '...in a mammalian cell under physiological conditions...each hour approximately 2,500 purine and 120 pyrimidine bases are detached, about 2,000 one-strand breaks occur, many cytosines are deaminated and at least 100 guanines are methylated'. Other authors suggest similar figures^{56–58}, and other kinds of DNA damage are also possible. To prevent mutation, all this damage must be correctly repaired at some cost.

The fidelity of DNA replication also requires many complicated mechanisms because a simple conformational correspondence between bases forming Watson-Crick pairs cannot pro-

vide good selectivity⁵⁹. In *Escherichia coli*, DNA-polymerase detaches a large proportion even of correctly attached nucleotides in the process of proof reading⁶⁰, one of the mechanisms for increasing fidelity^{61–63}. This and other mechanisms, including those acting immediately after replication^{64,65}, involve costs in both excessive energy requirements and time delay^{66,67}. With fixed accuracies of the separate stages (limited by chemical factors), the attempt to achieve the zero error frequency through higher activity of controlling processes leads to the infinite growth of the cost⁶⁸. So, the correct question is not 'why mutations occur' but 'what mechanisms reduce the rate and at what cost'.

A population of *D. melanogaster* exposed to mutagenic conditions (radiation) for many generations has a decreased rate of spontaneous mutation without radiation⁶⁹. This indicates that the natural mutation rate exceeds the lowest possible value and it could be decreased further under intensified selection for its reduction. I suggest that the high cost of fidelity precludes this decrease in nature. So, although there are no quantitative data for the cost of fidelity in terms of individual fitness, the idea that it precludes the genome degradation rate becoming less than about 0.5 seems reasonable.

Data on mutation rates

Classic methods that took into account only mutations with drastic phenotype and/or fitness effects suggested that mutations are very rare events. Traditional figures for mutability per locus^{70–72} are about 10^{-6} – 10^{-7} . The overall numbers of functional units (loci) are about 10^4 in the *Drosophila* genome and perhaps 10^5 in *Homo*. Hence, this implies low ($\ll 1$) mutation rates even for the whole genome of mammals. Mutations with only small effects are, however, much more frequent, as only a minority of possible genotype alterations leads to a qualitative change of function³⁴. To study such mutations one must observe changes at the molecular level, namely in protein or (preferably) DNA sequences. The necessary methodology of protein electrophoresis and DNA hybridization or (later) sequencing became available in the late 1960s. Alternatively, it is possible to count mutations by their effects on fitness if the method allows the detection of small changes. Two approaches to mutation rate measuring are possible, the *direct* comparison of parental and progeny genotypes and the *indirect* estimation from data on population parameters (variability) at a fixed time. The direct approach yields two sets of data:

Short-term mutation accumulation. Observations may last one generation (parent-offspring genotype comparison) or several. As in this case the mutation rate per locus is low, an analysis of large samples is necessary. Protein electrophoresis is appropriate, but suitable methods at the DNA level are still absent. Fitness monitoring is also applicable, as it counts mutations in the whole genome.

Long-term measurements. Although ancestors' genotypes are unavailable, one can compare genotypes of different progenies of a common ancestor. In this case mutations are frequent enough to allow the use of DNA analysis methods.

In both cases it is necessary to exclude selection in the course of accumulation of mutations. As most non-neutral mutations are deleterious^{73,74}, selection usually leads to underestimation of mutation rates. If some DNA region is really neutral (silent, excessive, junk and so forth), it accumulates mutations at the rate of their appearance in over both short and long time-scales^{73,75}. As most methods yield data per locus or per nucleotide, it is necessary to recall that minimal haploid genome sizes are about 3×10^7 for algae and fungi, 10^8 for lower invertebrates and higher plants, 3×10^8 for echinoderms and lower vertebrates and 3×10^9 for mammals^{76,77}.

Here I present a brief review of recent results. The most extensive data on human parent-offspring genotype comparisons have been obtained for the Japanese population^{78,79}. Three mutations altering the electrophoretic mobility of some

blood enzymes were observed in about 500,000 locus tests. Some substitutions lead to enzyme inactivation, but many substitutions are synonymous, that is, do not cause any protein change. As electrophoresis reveals that about half of all substitutions do not cause protein inactivation, the overall substitution rate is about 2×10^{-5} per locus or 2×10^{-8} per nucleotide, or about 100 per diploid genome per generation. Analogous data for *D. melanogaster*⁸⁰ revealed a similar frequency of noninactivating substitutions per nucleotide, which implies about 2 substitutions per diploid genome. In this case mutations to inactive alleles were also counted and their rate was even higher. Some other data also indicate that insertions and deletions, which normally inactivate protein when they occur in the coding region, are at least as frequent as substitutions⁸¹⁻⁸⁴.

Measurement of mutation rate through viability-change monitoring during several generations in *D. melanogaster*^{36,85} showed about one mutation lowering larval competitive ability occurred in the diploid genome in each generation, on the average. Probably the overall mutation rate is much higher because (1) neutral and nearly neutral mutations were neglected; (2) deleterious mutations influencing other fitness components (for example, those active only at the imago stage and responsible for fertility, mating success and/or longevity) were not counted; and (3) competitive ability under simplified experimental conditions, where factors of selection such as starvation and cold were excluded, probably revealed mutations only in a subset of genes active at the larva stage. Data on quantitative traits other than fitness also demonstrate measurable mutation rates but it is impossible to express them in terms of genome mutation rate⁸⁶.

Data on long-term mutation accumulation have become abundant in recent years. Comparison of pseudogene sequences reveal evolution rates (and probably mutation rates) of about^{87,88} 1×10^{-8} or even⁸⁹ 1×10^{-7} per nucleotide per generation in mammals. Data on DNA hybridization also show similar evolution rates⁹⁰⁻⁹². Probably the average genomic rate of evolution is lower than mutation rate (because of selection) and the rate of change of the most rapidly evolving fraction⁹³⁻⁹⁴ is more relevant. Evolution rate of DNA in some invertebrates is higher than that of mammals^{90,93}. This should be compared with the indirect mutation rate, estimated by various methods⁹⁵, which yield human population values of about 1×10^{-8} per nucleotide for the noninactivating substitution rate⁹⁶.

These approaches give remarkably similar results and suggest that classical figures for genomic mutation rates are underestimated by several orders of magnitude. The most abundant data for mammals indicate mutation rates at about 100 per diploid genome per generation. Scarce data on invertebrates suggest about 10 mutations per genome. Unfortunately there are no comparable data for plants and fungi where patterns of reproduction are highly varied. We suggest that the deleterious mutation rate per genome is about 1 in these taxa, which allows maximal diversity of the reproduction modes.

As we are interested in the rate of deleterious mutations, the problem of what part of the genome is functional becomes the most important. Some authors, starting from high genome mutation-rate data, claim that most of the genome is functionless because a high deleterious mutation rate would otherwise lead to an excessive mutation load⁹⁰. As shown above, if we assume truncation selection this conclusion is not valid, so the problem remains. Unfortunately, here we have only indirect evidence.

In protein-coding regions insertions and deletions are practically always deleterious, as are about 9/10 of nonsynonymous⁷³ and a smaller proportion of synonymous^{73,88} nucleotide substitutions. But although about half of the unique DNA (which makes up the majority of small genomes) is transcribed in some tissues, only 2-5% of it codes for proteins^{76,77,97}. Recent data reveal many functional regions in noncoding DNA, such as promoters, enhancers and other transcription regulators, and signals for splicing, replication and recombination⁹⁸. Both direct sequence

comparison^{84,89,99} and molecular hybridization^{90,94} indicate that many sequences evolve more slowly than pseudogenes, which implies some selection. Other evidence also suggests that a large proportion of the genome is functioning^{100,101} but the data are insufficient for quantitative conclusions. For all these reasons, I believe it is unlikely that mutations at 99% of the genome are neutral, which would be necessary to make the rate of deleterious mutations in mammals acceptable from the traditional point of view. As Neel *et al.*⁷⁸ wrote, 'The amount of selfish DNA is steadily shrinking. The question of how our species accommodates such mutation rates is central to evolutionary thought'.

As we have seen, truncation selection with large average genome contamination can solve this problem. With $u \gg 1$, reproduction must be sexual and selection must be truncation-like or the mutation load will be too large. Thus, if the values of $u > 1$ are justified, they strongly support the validity of the deterministic mutation concept of evolution of sexual versus obligate asexual reproduction.

Selection against mutations

The problem of the evolutionary processes in sexual populations remains, however, because here v , and not u , is decisive, and v can easily be small even when u is large, providing that $c > u^2$ (see above). It is very difficult to measure v directly as existing methods do not permit evaluation of the genome contamination of an individual. A reasonable approach is to measure the intensity of selection against mutations. My estimate of $0.5 < v < 2.0$ implies strong selection, as the mutation load even with $v = 0.5$ is at least ~50%. Fitness variance should also be large in this case, as well as average selection against new mutation, which is about $1/u$ (ref. 33), although average selection against mutation existing at equilibrium may be lower as the population is enriched by mutations with slight effects.

Unfortunately, there are no data on natural populations to confirm or reject the reality of this intensity of selection, because measuring of selection in nature is very difficult¹⁰²⁻¹⁰⁴. In some cases strong selection has been reported¹⁰⁵ but it is difficult to separate selection against mutations from other forms of selection. In all species, each female produces an average of about 10 progeny during her reproductive life, so selection with the required intensity is possible. Data on controlled populations are hardly relevant, as experimental conditions are quite different from natural ones and it may take a long time for a new equilibrium to be established¹⁰⁶. Studying correlations between the fitnesses of relatives¹⁰⁷ is a reasonable approach but the problem of measuring both fitness and the degree of relatedness in nature remains. In a human population not exercising any birth control, a high correlation has been reported between the fertility of relatives, which is consistent with strong selection¹⁰⁸.

The shape of the selection curve is also very important, because most previous results depend on the assumption that selection is similar to truncation. This type of selection is reasonable from a theoretical point of view, for example, if only the winners in pair competitions reproduce. Some data also indicate this type of selection in *Drosophila*¹⁰⁹ and RNA viruses¹¹⁰ but more data are necessary.

Applications to human genetics

The idea is well established that all normal human beings carry a number of slightly deleterious mutations and some purifying selection is unavoidable to keep a population in equilibrium³⁴. The hypothesis presented here, if correct, leads to some new quantitative conclusions about mutation-selection balance. We have suggested that $u \gg 1$ and v is about 1. Suppose that at some moment selection against mutations with slight effects disappears. If such conditions remain stable for a long time, the average genome contamination will be doubled after u gener-

ations, as without selection in each generation the average genome contamination increases by u , and the whole increase needed to double the initial contamination equals u^2 (when v is assumed to be equal to 1). If $u = 10$, the contamination will be doubled after 10 generations; with $u = 100$, it takes 100 generations. This seeming paradox arises because the starting contamination value in the former case is about 100 mutations per genome, whereas in the latter case it is about 10,000 mutations. Although some selection certainly occurs even in civilized human populations¹¹¹⁻¹¹³, it probably eliminates mainly severe genetic damage, as mildly affected individuals can reproduce successfully in a supportive environment.

It is not clear at what point the increase in the number of slightly deleterious mutations will begin to produce phenotypes

maladapted even under good conditions; in other words, when does hard selection begin to operate? It is possible that the physiology of humans is designed to tolerate only natural contamination and even a relatively small rise would lead to drastic effects. The suggested increase in contamination is now probably being masked at the phenotype level by the overall improvement in living conditions. Such an increase is undesirable as it is practically irreversible. But before considering the possible methods of preventing¹¹⁴ an increase in genome contamination, it is necessary to obtain more data on the parameters of the mutation process and selection against them.

I thank J. F. Crow, B. Charlesworth and J. Maynard Smith for valuable suggestions, although the responsibility for opinions expressed is my own.

- Levin, B. R. in *The Evolution of Sex: an Examination of Current Ideas* (eds Michod, R. E. & Lewin, B. R.) 194-211 (Sinauer, Sunderland, 1988).
- Weismann, A. *Nature* **36**, 607-609 (1887).
- Lloyd, D. G. *Evol. Biol.* **13**, 69-91 (1980).
- Muller, H. J. in *Genetics, Paleontology and Evolution* (eds Jepsen, G. L., Mayr, E. & Simpson, G. G.) 421-447 (Princeton University Press, 1949).
- Maynard Smith, J. *The Evolution of Sex* (Cambridge University Press, 1978).
- Bell, G. *The Masterpiece of Nature* (University of California Press, San Francisco, 1982).
- Charlesworth, B. *J. theor. Biol.* **84**, 655-672 (1980).
- Uenoyama, M. K. *Evolution* **40**, 388-404 (1986).
- Kondrashov, A. S. *Genetics* **111**, 635-653 (1985).
- Crow, J. F. in *The Evolution of Sex: an Examination of Current Ideas* (eds Michod, R. E. & Lewin, B. R.) 56-73 (Sinauer, Sunderland, 1988).
- Maynard Smith, J. in *The Evolution of Sex: an Examination of Current Ideas* (eds Michod, R. E. & Lewin, B. R.) 106-125 (Sinauer, Sunderland, 1988).
- Bernstein, H., Hopf, F. A. & Michod, R. E. in *The Evolution of Sex: an Examination of Current Ideas* (eds Michod, R. E. & Lewin, B. R.) 139-160 (Sinauer, Sunderland, 1988).
- Bengtsson, B. O. *Genet. Res.* **47**, 771-780 (1986).
- Kelley, S. E., Antonovics, J. & Schmitt, J. *Nature* **331**, 714-716 (1988).
- Harshman, L. G. & Futuyma, D. J. *Am. Nat.* **125**, 896-902 (1985).
- Griffin, M. J. *theor. Biol.* **102**, 447-458 (1983).
- Maynard Smith, J. *Genet. Res.* **51**, 59-64 (1988).
- Case, T. J. & Bender, E. A. *J. theor. Biol.* **90**, 181-190 (1981).
- Moore, W. S. & Hines, W. G. S. *J. theor. Biol.* **92**, 301-316 (1981).
- Schmidt, J. & Ehrhardt, D. W. *Evolution* **41**, 579-590 (1987).
- Willson, M. F. *et al. Am. Nat.* **129**, 304-311 (1987).
- Muller, H. L. *Mutat. Res.* **1**, 2-9 (1964).
- Manning, J. T. *Hereditas* **50**, 15-19 (1983).
- Shields, W. M. *Philopatry, Inbreeding and the Evolution of Sex* (State University of New York Press, Albany, 1982).
- Pamilo, P., Nei, M. & Li, W.-H. *Genet. Res.* **49**, 135-146 (1987).
- Penny, D. *Biol. J. Linn. Soc. Lond.* **25**, 209-220 (1985).
- Reaney, D. C. *Trends Genet.* **2**, 41-46 (1986).
- Kimura, M. & Maruyama, T. *Genetics* **54**, 1337-1351 (1966).
- Crow, J. F. in *Mathematical Topics in Population Genetics* (ed. Kojima, K.) 128-177 (Springer, Berlin, 1970).
- Kimura, M. & Crow, J. F. *Proc. natn. Acad. Sci. U.S.A.* **75**, 6168-6171 (1978).
- Crow, J. F. & Kimura, M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 396-399 (1979).
- Kondrashov, A. S. *Genet. Res.* **40**, 325-332 (1982).
- Kondrashov, A. S. *Genet. Res.* **44**, 199-217 (1984).
- Muller, H. J. *Am. J. hum. Genet.* **2**, 111-176 (1950).
- Crow, J. F. *Evolution* **37**, 863-865 (1983).
- Crow, J. F. & Simmons, M. J. in *The Genetics and biology of Drosophila* Vol. 3c (eds Ashburner, M., Carson, H. L. & Thompson, N. J. Jr) 1-35 (Academic, New York, 1983).
- Dickson, D. P. E. & Manning, J. T. *Hereditas* **53**, 717-720 (1984).
- Manning, J. T. & Dickson, D. P. E. *Acta biotheoret.* **35**, 149-169 (1986).
- Brooks, L. D. in *The Evolution of Sex: an Examination of Current Ideas* (eds Michod, R. E. & Lewin, B. R.) 87-105 (Sinauer, Sunderland, 1988).
- Jain, S. K. *Genetics* **66**, 21-27, 1985.
- Taylor, P. D. & Williams, G. C. *Theor. Populat. Biol.* **22**, 392-409 (1982).
- Charlesworth, D. & Charlesworth, B. A. *Rev. Ecol. Syst.* **18**, 237-268 (1987).
- Cvetkovic, D. & Tuccia, N. Z. *Zool. Syst. Evol.* **24**, 298-306 (1986).
- Heisler, I. L. *Evolution* **38**, 1283-1295 (1984).
- Kondrashov, A. S. *J. theor. Biol.* **131**, 487-496 (1988).
- Rice, W. R. *Evolution* **42**, 817-820 (1988).
- Feldman, M. W., Christiansen, F. B. & Brooks, L. D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4838-4841 (1980).
- Redfield, R. J. *Genetics* **119**, 213-221 (1988).
- Kondrashov, A. S. *Hereditas* **52**, 307-308 (1984).
- Burt, A. & Bell, G. *Nature* **326**, 803-805 (1987).
- Sturtevant, A. H. *Q. Rev. Biol.* **12**, 464-467 (1937).
- Drake, J. W., Glikman, B. W. & Ripley, L. S. *Am. Scient.* **71**, 621-630 (1983).
- Sargentini, N. J. & Smith, K. C. *Mutat. Res.* **154**, 1-27 (1984).
- Vilenchik, M. M. *DNA Instability and Delayed Consequences of Irradiation* (Energoatomizdat, Moscow, 1987) (in Russian).
- Vilenchik, M. M. *The Biological Foundations of Senescence and Longevity* (Znanie, Moscow, 1987) (in Russian).
- Mazin, A. L., Gimadudinov, O. A., Turkin, S. I., Burtseva, N. N. & Vaniushin, B. F. *Molekularnaya Biologiya* **19**, 903-914 (1985).
- Kunkel, T. A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1494-1498 (1984).
- Loeb, L. A. & Preston, B. D. A. *Rev. Genet.* **20**, 201-230 (1986).
- Chuprina, V. P. & Poltev, V. I. *Nucleic Acids Res.* **13**, 141-154 (1985).
- Fersht, A. R., Krill-Jones, J. W. & Tsui, W.-C. *J. molec. Biol.* **156**, 37-51 (1982).
- Kaguni, L. S., DiFrancesco, R. A. & Lehman, I. R. *J. biol. Chem.* **259**, 9314-9319 (1984).
- Krutakov, V. M., Belyakova, N. V., Kleiner, N. E., Legina, O. K. & Shevelev, I. V. *Doklady Akademii Nauk SSSR* **272**, 1491-1494 (1983).
- Kraevski, L. A. & Kukhanova, M. K. *Results in Science and Technology (VINITI). Ser. Molec. Biol.* **22**, 100-150 (1986) (in Russian).
- Lindahl, T. A. *Rev. Biochem.* **51**, 61-87 (1982).
- Carraway, M., Youderian, P. & Marinus, M. G. *Genetics* **116**, 343-347 (1987).
- Blomberg, C. J. *theor. Biol.* **128**, 87-107 (1987).
- Blomberg, C. J. *theor. Biol.* **115**, 241-268 (1985).
- Kirkwood, T. B. L., Rosenberger, R. F. & Galas, D. J. (eds) *Accuracy in Molecular Processes: its Control and Relevance to Living Systems* (Chapman & Hall, London, 1986).
- Nothel, H. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1045-1048 (1987).
- Neel, J. V. *J. Hered.* **74**, 2-15 (1983).
- Crow, J. F. & Denniston, C. in *Advances in Human Genet.* Vol. 14 (eds Harris, H. & Hirschhorn, K.) 59-123 (Plenum, New York, 1985).
- Vogel, F. & Motulsky, A. G. *Human Genetics* (Springer, New York, 1986).
- Kimura, M. *The Neutral Theory of Molecular Evolution* (Cambridge University Press, 1983).
- Watterson, G. A. *Genet. Res.* **50**, 155-164 (1978).
- Li, W.-H., Gojobori, T. & Nei, M. *Nature* **292**, 237-239 (1981).
- Lewin, B. *Genes* (Wiley, New York, 1985).
- Gazaryan, K. G. & Tarantul, V. Z. *Eukaryotic Genome* (Moscow University Press, 1983) (in Russian).
- Neel, J. V. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 389-393 (1986).
- Neel, J. V. *et al. Am. J. hum. Genet.* **42**, 663-676 (1988).
- Mukai, T. & Cockerham, C. C. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2514-2517 (1977).
- Fukasawa, K. M. *et al. Genetics* **115**, 177-184 (1987).
- Piatigorsky, J., Norman, B. & Jons, R. E. *J. molec. Evol.* **25**, 308-317 (1987).
- Ashman, C. R. & Davidson, R. L. *Proc. natn. Acad. Sci. U.S.A.* **84**, 3354-3358 (1987).
- Kreitman, M. *Nature* **304**, 412-416 (1983).
- Mukai, T., Chigusa, S. T., Mettler, L. E. & Crow, J. F. *Genetics* **72**, 335-355 (1972).
- Lynch, M. *Genet. Res.* **51**, 137-148 (1988).
- Li, W.-H., Tanimura, M. & Sharp, P. M. *J. molec. Evol.* **25**, 330-342 (1987).
- Li, W.-H. & Tanimura, M. *Nature* **326**, 93-96 (1987).
- Mermer, B., Colb, M. & Krontiris, T. G. *Proc. Natn. Acad. Sci. U.S.A.* **84**, 3320-3324 (1987).
- Britten, R. J. *Science* **231**, 1393-1398 (1986).
- Catzeffis, F. M., Sheldon, F. H., Ahlquist, J. E. & Sibley, C. B. *Molec. Biol. Evol.* **4**, 242-253 (1987).
- Powell, J. R., Caccione, A., Amato, G. D. & Yoon, C. *Proc. natn. Acad. Sci. U.S.A.* **83**, 9090-9093 (1986).
- Zwibel, L. J., Cohn, V. H., Wright, D. R. & Moore, G. P. *J. molec. Evol.* **19**, 62-71 (1982).
- Roberts, J. W. *et al. J. molec. Evol.* **22**, 99-107 (1985).
- Pollak, E. *Am. J. hum. Genet.* **38**, 209-227 (1986).
- Neel, J. V., Mohrenweiser, H. W., Rothman, E. D. & Naidu, J. M. *Am. J. Hum. Genet.* **38**, 649-666 (1986).
- Bossi, B., Hall, L. M. C. & Spierer, P. *EMBO J.* **3**, 2537-2541 (1984).
- Nussinov, R., Owens, J. & Maizel, J. V. *Biochem. biophys. Acta* **866**, 109-119 (1986).
- Kreitman, M. & Aguade, M. *Genetics* **114**, 93-110 (1986).
- Mouchiroud, D., Gautier, C. & Bernardi, G. *J. molec. Evol.* **27**, 311-320 (1988).
- Schaffner, G., Schrin, S., Miller-Baden, B., Weber, F. & Schaffner, W. *J. molec. Biol.* **201**, 81-90 (1988).
- Endler, J. A. *Natural Selection in the Wild* (Princeton University Press, 1986).
- Mitchell-Olds, T. & Shaw, R. G. *Evolution* **41**, 1149-1161 (1987).
- Golding, G. B. *Genet. Res.* **49**, 71-82 (1987).
- Taylor, C. E. & Kekic, V. *Evolution* **42**, 197-199 (1988).
- Mackay, T. F. C. *Genet. Res.* **47**, 59-70 (1986).
- Crow, J. F. & Denniston, C. *Science* **212**, 888-893 (1981).
- Bol'shakova, L. P. & Revazov, A. A. *Genetika* **24**, 340-349 (1988).
- Mukai, T. *Genetics* **61**, 749-761 (1969).
- Chao, L. J. *theor. Biol.* **133**, 99-112 (1988).
- Czeizel, A. & Sankaranarayanan, K. *Mutat. Res.* **128**, 73-103 (1984).
- Czeizel, A. & Tusnady, G. *Aetiological Studies of Isolated Common Congenital Abnormalities in Hungary* (Akademiai Kiado, Budapest, 1984).
- Wilcox, A. J. *et al. New Eng. J. Med.* **319**, 189-199 (1988).
- Alekperov, U. K. *Antimutagenesis* (Nauka, Moscow, 1984) (in Russian).

Are global cloud albedo and climate controlled by marine phytoplankton?

Stephen E. Schwartz

Environmental Chemistry Division, Brookhaven National Laboratory, Upton, New York 11973, USA

The recent suggestion that dimethylsulphide emissions by marine phytoplankton control global albedo and mean temperature would also imply a strong climatic influence of man-made SO₂. Anthropogenic SO₂ emissions exceed marine emissions of dimethylsulphide globally and are confined largely to the Northern Hemisphere. But no such influence of SO₂ emissions is found either in the present cloud component of planetary albedo or in 100-year temperature records.

RECENTLY, Charlson, Lovelock, Andreae and Warren¹ (hereinafter CLAW) proposed that production of dimethylsulphide (DMS) by marine microorganisms is the major source of aerosol sulphate in the remote marine atmosphere and, in turn, of cloud condensation nuclei (CCN). They suggest that this process, by influencing planetary albedo, comprises a component of a biological mechanism for climate regulation. Specifically, it is hypothesized that an increase in marine DMS emissions would increase the number density of aerosol sulphate particles and, in turn, the number density of cloud droplets, thus enhancing cloud albedo. This enhancement would reduce global temperature generally, including ocean temperatures, and would consequently reduce marine productivity and DMS emission; that is to say, the process is a negative feedback system for control of the Earth's temperature. Such a feedback system was subsequently and independently suggested by Mészáros².

Here I examine the hypothesized link between gaseous precursors of aerosol sulphate, and global albedo and mean temperature. Because marine DMS emissions are substantially lower than present emissions of other gaseous precursors of atmospheric sulphate (principally SO₂ associated with fossil fuel combustion) globally and especially in the Northern Hemisphere, then consequently, if, as CLAW suggest, marine DMS emissions exert a major influence on global mean albedo and temperature, such an influence should also result from anthropogenic SO₂ emissions. The potential for such anthropogenic influence on cloud albedo and global mean temperature has been noted previously^{3,4}. But CLAW explicitly "ignore the perturbations of the atmospheric sulphur cycle by man-made fluxes of SO₂... and... consider only the natural fluxes, which currently represent about 50% of the total flux of gaseous sulphur to the atmosphere, and which still dominate the atmospheric sulphur cycle in the Southern Hemisphere". I argue here that the increase in SO₂ emissions over the past 100 years and the contrast in

gaseous sulphur emissions between the Northern Hemisphere (NH) and Southern Hemisphere (SH) constitute a global experiment that tests the hypothesis of ref. 1, that planetary albedo and mean temperature are regulated by marine DMS emissions.

It will be seen that a comparison of the present NH and SH cloud component of planetary albedo and of 100-year temperature records yields no indication of any influence of anthropogenic SO₂ emissions that is either qualitatively or quantitatively consistent with the expectations in this respect based on the CLAW hypothesis.

Emission rates

Emissions of reduced sulphur compounds and of sulphur dioxide have been critically reviewed recently⁵⁻⁷ and are summarized in Table 1. Andreae⁶ estimates total global biogenic emissions of sulphur gases to be 95 ± 30 Tg of sulphur per year, DMS and H₂S each accounting for about half of this total (43 ± 20 and 32 ± 27 Tg S yr⁻¹, respectively). The estimate⁶ for marine DMS emissions, 40 ± 20 Tg S yr⁻¹, is based on concentrations of DMS in ocean surface water at a variety of locations together with estimates of sea-air exchange rates. A lower DMS emission rate of 16 Tg S yr⁻¹ (uncertain to a factor of two) is given by Bates *et al.*⁸. Natural emission rates in Table 1 are consistent with budget estimates based on considerations of turnover times and of concentrations of sulphur compounds in the atmosphere⁹.

On a global basis, SO₂ emissions are roughly equal to biogenic emissions and are at least twice as great as marine DMS emissions. A *prima facie* case exists, therefore, that anthropogenic sulphur emissions cannot be neglected in examining the possible influence of biogenic gaseous sulphur compounds on global cloud albedo and climate. Also, as anthropogenic SO₂ emissions have increased to their present level almost entirely within the last 100 years (ref. 5 and H. Rodhe, private communication cited in ref. 7), any climatic influence of these emissions should be discernible in records over this period¹⁰. Furthermore, the asymmetric distribution of anthropogenic SO₂ emissions over the globe suggests that the influence of gaseous sulphur emissions on climate may also be tested by interhemispheric comparisons, significant interhemispheric transport (on a timescale of one to two years) being precluded by the short atmospheric residence times of these substances^{9,11}.

Distribution of aerosol sulphate

For sulphate derived from anthropogenic SO₂ to serve as a surrogate to test the CLAW hypothesis, this material must be widespread throughout the NH. The geographical distribution of sulphate attributable to anthropogenic emissions of SO₂ is examined in Table 2, which presents concentrations of aerosol sulphate in boundary-layer air at locations remote from

Table 1 Gaseous sulphur emission rates

Emission source	Northern Hemisphere	Southern Hemisphere	Global
Marine DMS	17	23	40
Total marine biogenic	22	28	50
Terrestrial biogenic	32	16	48
Total biogenic	54	44	98
Anthropogenic SO ₂	98	6	104
Total	150	50	200

Data are from Cullis and Hirschler⁵, except for marine DMS (Andreae⁶), apportioned according to ocean surface area: 42% NH and 58% SH. Emission rates are in units of Tg S yr⁻¹.

Table 2 Measured concentration of sulphate (or sulphur) aerosol at remote locations in the Northern and Southern Hemisphere

Location	Sulphate concentration* (ng S m ⁻³)	Comments	Ref.
Northern Hemisphere			
<i>High-latitude sites</i>			
Alert, Mould Bay, Igloolik; Canadian Arctic (66–83° N)	(200–1,000) (20–70)	NSS SO ₄ ²⁻ ; 1-week samples; 3- to 4-year data record Winter–Spring	15
Faeroe Islands (62° N)	1,100 (700–1,400) 140 (30–230)	Summer NSS SO ₄ ²⁻ British trajectory (~1,000 km); 4 1-day samples	13
Velen, Sweden (58° N)	960 [1.4] 60 [2.2]	Atlantic trajectory; 4 1-day samples Sub-μm S; 1-day samples British trajectory (~1,000 km); 12 samples North Sea trajectory (Northwest air); 14 samples	18
<i>Atlantic and Caribbean</i>			
Western North Atlantic (33–38° N, 65–70° W)	800 ± 500	Sub-μm S; 26 8- to 68-hour samples	18
Bermuda (32° N)	530 [3.0]	Sub-μm S; 10 2- to 4-day samples	18
Bermuda	1,300 630 390 330 650 ± 230	Sub-2.5-μm NSS SO ₄ ²⁻ ; 39 1-day samples Northeast US trajectory (~1,200 km) Southeast US trajectory (~1,100–1,500 km) Caribbean trajectory Southeast trajectory All	14
Barbados (13° N, 60° W)	300 ± 190 120 ± 65 250 ± 180	NSS SO ₄ ²⁻ ; 1-day samples 'High dust'—Europe or Africa trajectory 'Low dust' All	17
<i>Pacific</i>			
Eastern Pacific Ocean off Washington State (47–48° N)	140 ± 60 80 ± 30	Aircraft sampling, boundary layer; 3 flights; Pacific trajectories; May Total NSS SO ₄ ²⁻ Sub-1.5-μm NSS SO ₄ ²⁻	20
Midway Island (28° N)	260 ± 90 100 ± 30 190 ± 260	NSS SO ₄ ²⁻ ; 1-week samples; onshore flow only 'Dusty'; 29 samples 'Clean'; 27 samples All; 56 samples	16
Oahu (21° N)	160 ± 130 90 ± 80 120 ± 110	NSS SO ₄ ²⁻ ; 1-week samples; onshore flow only 'Dusty'; 24 samples 'Clean'; 32 samples All; 56 samples	16
Guam (17° N)	150 ± 240	NSS SO ₄ ²⁻ ; 1-week samples; onshore flow only; 49 samples	16
Belau (7° N)	210 ± 190	NSS SO ₄ ²⁻ ; 1-week samples; onshore flow only; 40 samples	16
Fanning Island (4° N)	210 ± 50	NSS SO ₄ ²⁻ ; 1-week samples; onshore flow only; 48 samples	16
Southern Hemisphere			
Samoa (14° S)	60 [2.4]	Sub-μm S; 17 3- to 5-day samples	18
Samoa	130 ± 50	NSS SO ₄ ²⁻ ; 42 1-week samples	19
New Caledonia (22° S)	170 ± 130	NSS SO ₄ ²⁻ ; 46 1-week samples	19
Norfolk Island (29° S)	110 ± 50	NSS SO ₄ ²⁻ ; 41 1-week samples	19
Cape Grim, Tasmania (41° S)	90 ± 20	NSS SO ₄ ²⁻ ; multi-year data record under 'baseline' conditions	43
Tasmania, off West Coast	(22–70)	NSS SO ₄ ²⁻ ; aircraft sampling, boundary layer; 8 flights; ocean trajectories; Austral summer	21
Punta Arenas, Chile (54° S)	52 [0.3] 83 11	Austral summer; sub-μm S; 9 3- to 5-day samples Austral summer; NSS S; 140 4-hour samples Austral winter; sub-μm S; 1 5-day sample	18
South Pole	76 ± 24 29 ± 10	Austral summer; total S; sea salt small; 40 several-day samples Austral winter; total S; sea salt 4–10%; 49 several-day samples	44

* Mean ± standard deviation or geometric mean [geometric standard deviation] or (range).

† NSS SO₄²⁻: Non-seasalt sulphate.

industrial activity. Such locations are pertinent because aerosol sulphate concentrations are obviously high in industrialized regions and at locations immediately influenced by transport from these regions. By comparison, median 24-hour boundary-layer sulphate concentrations at non-urban locations of the north-eastern United States are ~2,000 ng S m⁻³ (ref. 12). Total aerosol samples are corrected for sea salt by reference to Na or Mg; this correction is generally unimportant for measurements

of only sub-micrometre (sub-μm) sulphate (that fraction arising from gas-to-particle conversion, which is of concern here). The data in Table 2 represent measurements by numerous investigators using a variety of techniques and protocols for widely differing sampling periods, so that any comparison of data from different investigators must be made with caution. Nonetheless, Table 2 indicates a pattern in which aerosol sulphate concentrations at remote NH locations substantially exceed those at

remote SH locations. Concentrations in the NH are mostly greater than $\sim 150 \text{ ng S m}^{-3}$, often much greater, whereas concentrations at remote locations in the SH are almost invariably less. Concentrations in the NH also exhibit substantially greater variability than those in the SH, indicating the intermittent presence, in the former case, of aerosol sulphate transported from distant regions.

For several NH sites, high concentrations have explicitly been ascribed to long-range transport from continental regions or, more specifically, from regions of industrial activity, by the trajectory of the air arriving at the site or by the presence of other air constituents. For example, high sulphate concentrations at the Faeroe Islands are attributed to transport from the United Kingdom, 1,000 km distant¹³, and high sulphate concentrations at Bermuda are attributed to transport from North America (1,100 km)¹⁴. The marked seasonality of sulphate concentrations in the Canadian Arctic¹⁵ is attributed to transport from Eurasian sources during late winter and spring; this seasonal cycle is exhibited also by a number of other anthropogenic constituents. High non-seasalt (NSS) sulphate at several North Pacific islands commonly coincides with soil dust, transported mostly from Asia¹⁶. Similarly, high NSS sulphate at Barbados, coincident with high incidence of soil dust, is attributed to transport from North Africa or Europe (distances of 5,000 km or more)¹⁷.

Several data sets in Table 2 afford direct comparison of measurements at NH and SH locations. Lawson and Winchester¹⁸ present measurements of sub- $\mu\text{m S}$ for which NH concentrations were an order of magnitude greater than SH concentrations. Prospero *et al.*¹⁶ and Saltzman *et al.*¹⁹ present an extensive set of measurements from several North Pacific¹⁶ and South Pacific¹⁹ islands; the mean NSS sulphate concentration for five NH islands was 30% greater than that for three SH islands.

Andreae *et al.*²⁰ and Berresheim *et al.*²¹ report aircraft measurements of NSS sulphate by similar techniques in boundary-layer air off the coast of Washington state²⁰ and off the west coast of Tasmania²¹, both sets of data relating to air behind cold fronts moving over open ocean without contact with land for several days. The NSS sulphate concentrations for Washington state ranged from 90 to 190 ng S m^{-3} , whereas those for Tasmania ranged from 20 to 70 ng S m^{-3} . NSS sulphate concentrations in the free troposphere at Tasmania (30–60 ng S m^{-3} at STP) were comparable to those in the boundary layer, but those at Washington state were substantially greater (150–270 ng S m^{-3} at STP). This situation argues strongly against a marine source. Indeed, Andreae *et al.*²⁰ suggest, in the case of the Washington state measurements, that the sulphur cycle in the free troposphere was dominated by transport from Asia.

Sulphate in precipitation

Table 3 presents volume-weighted mean concentrations of NSS sulphate obtained in long-term records of wet-only event sampling at several remote NH and SH sites²². The measurements indicate markedly greater concentrations in the NH, supporting the present assertion that sulphate derived from anthropogenic emissions exerts an influence on air composition throughout the NH.

Long-range transport of anthropogenic sulphur is indicated also by the dependence of the composition of precipitation on back trajectory. For samples collected in the eastern North Atlantic, Nyberg²³ found that for trajectories from North America (4,000–5,000 km), NSS SO_4^{2-} concentrations were 12–37 $\mu\text{eq l}^{-1}$, whereas for a trajectory emanating from the Azores, concentrations were $\sim 6 \mu\text{eq l}^{-1}$. Similarly, Jickells *et al.*²⁴ established that, in precipitation at Bermuda, systematically greater acidity was associated with trajectories from North America (mean H^+ concentration = 41 $\mu\text{eq l}^{-1}$) than with trajectories from the Caribbean (8 $\mu\text{eq l}^{-1}$) or the east (12 $\mu\text{eq l}^{-1}$); one can assume that the dependence of NSS SO_4^{2-} concentration on

trajectory was similar, because H^+ and NSS SO_4^{2-} were highly correlated (correlation coefficient $R = 0.94$; slope = 0.94).

Recently, Whelpdale *et al.*²⁵ concluded that even the most remote parts of the North Atlantic are undoubtedly affected by anthropogenic pollution and that only a few data are available in which precipitation is derived from air from 'clean' regions. Their estimate of sulphate concentrations in 'background' marine precipitation samples (6–8 $\mu\text{eq l}^{-1}$) substantially exceeds concentrations reported from remote SH sites (Table 3), indicat-

Table 3 Volume-weighted mean concentrations of non-seasalt sulphate in precipitation of remote sites

Location	NSS SO_4^{2-} $\mu\text{eq l}^{-1}$	Dates	Ref.
Poker Flat, Alaska (67° N)	10.2	5/80–5/81	22
Bermuda (32° N)	18.2	5/80–5/81	45
	13.9	8/82–5/84	46
San Carlos, Venezuela (2° N)	3.0	9/80–3/81	22
Katherine, Australia (14° S)	3.2	1980–1984	47
Amsterdam Island (38° S)	5.0	5/80–8/83	47
Torres del Paine, Chile (47° S)	2.1	1984–1985	47

ing that even this background contains a substantial component of anthropogenically derived sulphate. Considerably higher concentrations were reported for samples obtained with trajectories from North America or Europe.

Sulphate in polar ice

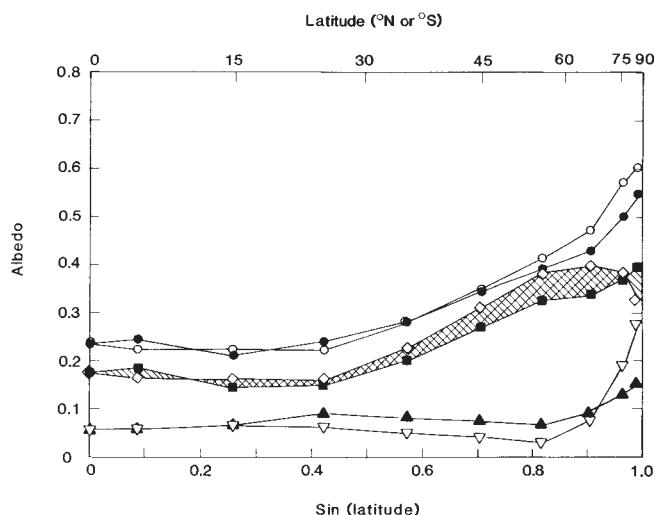
Substantial increases in sulphate concentrations over the past century have been noted in polar ice at NH sites. Concentration of NSS SO_4^{2-} in ice at Dye 3 Station, Greenland (65° N) has tripled since ~ 1900 ^{26,27}. Similarly, there has been a substantial increase in conductivity of melt water from core samples obtained at Ellesmere Island (81° N) from 1910 to 1980, which is attributed to increased deposition of sulphate, nitrate and associated cations²⁸. In contrast, a corresponding increase in sulphate concentration is not observed in ice cores obtained at Antarctica^{29,30}. One may conclude that deposition of NSS SO_4^{2-} at these remote NH sites has increased as a consequence of anthropogenic emissions and that the relative increase in this deposition is comparable to that in NH sulphur emissions.

Thus, comparisons of concentrations of NSS sulphate in aerosol, precipitation and ice cores at remote locations in the NH and SH establish that there are substantially greater concentrations throughout the NH than in remote SH locations, mainly as a result of anthropogenic emissions. These observations confirm my premise that if marine DMS emissions were to exert a control over global cloud albedo and mean temperature, then there should be an influence on these variables from anthropogenic SO_2 emissions.

Interhemispheric comparison of cloud albedo

A key component of the CLAW hypothesis is that the cloud contribution to planetary albedo should increase with increasing emissions of sulphate precursors. In the absence of measurements that would permit comparison of present and pre-industrial albedos, I seek to identify the effect of anthropogenic sulphur emissions on planetary albedo by a comparison of the present albedos of the two hemispheres. The expected magnitude of the interhemispheric albedo difference may be determined from the example given in ref. 1, which indicates that a 30% increase in CCN concentration would increase planetary albedo above marine clouds by 0.016, and average global albedo by 0.005, where the latter quantity takes into account the fraction of the Earth's surface covered by marine clouds. Following this example, and using a NH:SH CCN ratio of 3:1 (compare Table 1), I obtain a difference between average NH and SH albedos

Fig. 1 Average annual zonal mean over 10° latitude bands for Northern and Southern Hemispheres of total planetary albedo α_T and of cloud and clear-sky components, $F_C\alpha_C$ and $(1-F_C)\alpha_S$, respectively. The key is as follows: ●, total albedo (NH); ○, total albedo (SH); ■, Cloud component (NH); ◇, cloud component (SH); ▲, clear-sky component (NH); ▽, clear-sky component (SH). Single-hatching indicates that the NH cloud component exceeds the SH component; cross-hatching indicates the reverse. Equal distances on the lower abscissa correspond to equal areas on the Earth's surface.



of 0.023. Essentially the same result is obtained using the logarithmic dependence of albedo on CCN concentration proposed by Twomey *et al.*⁴. Although such an albedo difference is undoubtedly an overestimate, in view of the non-uniform distribution of sulphate concentrations in the NH, even a 50% difference in CCN concentrations between the two hemispheres (see Tables 2 and 3) would, in the absence of other factors, lead to an expected mean interhemispheric albedo difference of 0.008 (NH greater than SH).

In view of differences in land area and surface features of the two hemispheres (and resultant differences in surface albedo) I have looked for indication of an influence of NH SO₂ emissions by comparing not the total albedo of each hemisphere but the component of this albedo that is due to clouds, under the assumption that this quantity was more nearly comparable in the pre-industrial atmosphere. The cloud component of albedo is a direct measure of the contribution of clouds to solar radiative forcing of the Earth's heat budget³¹ and is thus an appropriate indicator of a possible influence of gaseous sulphur emissions on this budget.

Figure 1 shows, for each hemisphere, the annual-average mean of total albedo and of the cloud and clear-sky components of this albedo, over 10° latitudinal zones. The cloud component is evaluated³¹ as

$$F_C\alpha_C = \alpha_T - (1 - F_C)\alpha_S$$

where F_C represents average fractional cloud coverage³², α_C represents average cloud albedo, and α_T and α_S represent measured³³ average total and clear-sky albedo, respectively. There is no indication that the cloud component is systematically greater in the NH than in the SH; if anything, the reverse is true throughout the latitude range 15–75°, which range encompasses the bulk of the industrialized regions of the NH. These data thus give no evidence of the enhancement of the cloud component of mean NH albedo that would be expected if NH SO₂ emissions were to influence the Earth's heat budget and temperature in a manner consistent with the CLAW hypothesis.

Changes in temperature

A second test for the influence of anthropogenic SO₂ on climate is provided by examining temperature records for changes in global mean temperature or for hemispherical asymmetries. The example given by CLAW indicates that the hypothesized increase in mean global albedo of 0.005 associated with a 30% increase in CCN concentration results in a change in global average solar radiative heating, ΔQ , of -1.7 W m^{-2} . The resul-

tant equilibrium decrease in mean global temperature is $\Delta T = \Delta Q/\lambda$, where λ , the climate sensitivity parameter, is estimated³⁴ to be $1.8 \text{ W m}^{-2} \text{ K}^{-1}$ (uncertain to a factor of two); this yields $\Delta T = -1.0 \text{ K}$. (This value of ΔT , and others evaluated here, are all uncertain to a factor of two because of uncertainty in λ . Evidently CLAW used a somewhat lower value of λ , yielding $\Delta T = -1.3 \text{ K}$.) Analogously, an increase of 0.023 in mean NH albedo (which, according to the CLAW hypothesis, would result from a tripling of NH sulphur emissions) should give rise to a decrease in mean NH temperature of 4.4 K if, along with the increase in emissions and concentrations, the cooling effect were confined entirely to the NH^{35,36}. Alternatively, if the temperature change were distributed over both hemispheres, an average global temperature decrease of 2.2 K would be indicated. Again, because of spatial non-uniformities in CCN concentration, such temperature changes may be overestimates. But even a NH albedo increase of 0.008, which corresponds in the CLAW hypothesis to a uniform 50% increase in CCN concentration over pre-industrial values, would result in a NH temperature decrease of 1.5 K in the absence of interhemispheric coupling, or a global temperature decrease of 0.8 K if the cooling was distributed globally.

These estimates represent equilibrium temperature changes, whereas any currently observable temperature decrease corresponding to the indicated change in radiative forcing would be less, by a factor of about two, because of the lag time in reaching thermal equilibrium^{10,39}. It should be noted that such predicted temperature change is comparable to, and of opposite sign to, the change in global temperature that would be expected from increased concentrations of CO₂ and other 'greenhouse' gases^{10,34,37}. I return to this point later.

Several analyses of temperature records in the Northern and Southern Hemispheres over the past hundred years have recently been presented. Jones *et al.*^{38–40}, by examining records for land and sea-surface temperatures in the two hemispheres, present evidence for warming trends that are essentially identical in both hemispheres. The land temperatures record (1881–1984) indicates warming rates (K per century) of 0.51 (NH) and 0.47 (SH). The corresponding warming rates determined from the sea surface temperature record (1904–1984) are 0.60 (NH) and 0.57 (SH). Hansen and Lebedeff⁴¹ propose a global warming trend of $0.6 \pm 0.1 \text{ K per century}$, again with no discernible inter-hemispheric difference. Thus there is no cooling trend in either hemisphere, or any lesser warming in the NH relative to the SH, as a consequence of the increase in sulphur emissions in the NH over this time period, of magnitude such as would be expected according to the CLAW hypothesis.

Summary and discussion

The hypothesis that emissions of DMS by marine phytoplankton exert a regulatory influence on mean global cloud albedo and temperature has been tested by searching for a possible influence of anthropogenic SO_2 emissions on these properties. These emissions exceed estimated global marine DMS emission by a factor of two, are comparable to total biogenic gaseous sulphur emissions globally, are confined largely to the NH, and have reached their present levels within the past 100 years. Moreover, aerosol sulphate is widespread throughout the NH at concentrations well in excess of those in remote SH locations. These features of anthropogenic SO_2 emissions and resultant sulphate concentrations suggest that the influences of anthropogenic SO_2 emissions on mean hemispheric cloud albedo and on mean hemispheric or global temperature would be evident if these aspects of global climate were controlled by sulphate derived from marine DMS emissions. The absence of any excess mid-latitude cloud component of albedo in the NH relative to the SH, and the lack of any global cooling or of any reduced warming in the NH relative to the SH, indicate that these variables are not controlled by anthropogenic sulphates, and by extension, are not controlled by marine DMS emissions.

The apparent lack of observable influence of emissions of gaseous sulphur compounds on climate is puzzling. There is little doubt that SO_2 and reduced sulphur compounds are precursors to aerosol sulphate⁴². It is also well established that sulphate-containing aerosol particles are effective cloud condensation nuclei^{1,3}. There is also persuasive theoretical^{1,3,4} and observational³ evidence that the albedo of liquid-water clouds increases with cloud droplet number density. It thus seems reasonable that an influence on the cloud contribution to global mean albedo and temperature might be expected, as Twomey *et al.*⁴, CLAW¹ and Mészáros² have suggested. The observational evidence, however, suggests otherwise. Thus, unless there is a countervailing influence upon global or hemispheric cloud albedo or temperature that has masked the influence of SO_2 emissions, it seems unlikely that global cloud albedo and temperature are controlled by emissions of gaseous sulphur compounds.

One such influence that must be addressed is that of warming caused by CO_2 and other greenhouse gases. Current evaluations^{34,37} of expected global temperature changes since 1850 resulting from increased concentrations of CO_2 and other infrared-active gases (CH_4 , N_2O , chlorofluoromethanes) indicate an expected mean temperature increase of $\sim 1\text{ K}$. (This temperature change is evaluated from modelled radiative forcing^{34,37} with the same sensitivity parameter λ employed above and exhibits the same factor of two uncertainty.) Such an expected temperature increase is comparable to but somewhat larger than estimates of actual temperature changes^{38–41}, the difference being attributed^{10,37} both to the time lag in response of global temperature to the change in radiative forcing and to the uncertainty in λ , although in principle some of that difference might be due to enhanced cloud albedo from anthropogenic sulphate. The several greenhouse gases are, however, all long-lived relative to interhemispheric mixing time, and are hence uniformly mixed globally, leading to an essentially symmetric forcing function on temperature, whereas anthropogenic sulphate, confined largely to the NH, would exert its influence largely in that region. The lack of hemispherical asymmetry in the rate of change in global temperature^{38–41} thus suggests that compensation of cooling from enhanced cloud albedo by warming caused by greenhouse gases does not account for the observations. Also, greenhouse heating would not account for the lack of enhanced cloud contribution to NH albedo.

In conclusion, it would seem that the lack of discernible response of mean global or hemispheric albedo or temperature to anthropogenic SO_2 emissions indicates that control of these properties is too complex to be governed by a single variable of this type. Nonetheless, the potential for a substantial human influence on global climate by the mechanism examined here makes it mandatory to gain a thorough understanding of the processes that control cloud albedo and its influence on global climate.

I thank Peter Daum, Paul Falkowski, Martin Hoffert, Paul Michael, Leonard Newman and Douglas Wallace for valuable discussions. This work was supported by the National Acid Precipitation Assessment Program through the PRECP Program funded by the US Department of Energy.

Received 16 May; accepted 18 October 1988.

1. Charlson, R. J., Lovelock, J. E., Andreae, M. O. & Warren, S. G. *Nature* **326**, 655–661 (1987).
2. Mészáros, E. *Atmos. Envir.* **22**, 423–424 (1988).
3. Twomey, S. *Atmospheric Aerosols* 278–290 (Elsevier, New York, 1977).
4. Twomey, S. A., Piepgrass, M. & Wolfe, T. L. *Tellus* **36B**, 356–366 (1984).
5. Cullis, C. F. & Hirschler, M. M. *Atmos. Envir.* **14**, 1263–1278 (1980).
6. Andreae, M. O. in *The Biogeochemical Cycling of Sulfur and Nitrogen in the Remote Atmosphere* (eds Galloway, J. N., Charlson, R. J., Andreae, M. O. & Rodhe, H.) 5–25 (Reidel, Dordrecht, 1985).
7. Ryaboshapko, A. G. in *The Global Biogeochemical Sulfur Cycle SCOPE Report 19* (eds Ivanov, M. V. & Freney, J. R.) 203–296 (Wiley, New York, 1983).
8. Bates, T. S., Cline, J. D., Gammon, R. H. & Kelly-Hansen, S. R. *J. geophys. Res.* **92**, 2930–2938 (1987).
9. Rodhe, H. & Isaksen, I. *J. geophys. Res.* **85**, 7401–7409 (1980).
10. Wigley, T. M. L. & Schlesinger, M. E. *Nature* **315**, 649–652 (1985).
11. Schwartz, S. E. *Tellus* **31**, 530–547 (1979).
12. Hidy, G. M. & Mueller, P. K. *The Sulfate Regional Experiment Report EA 1901*, Vol. 2 (Electric Power Research Institute, Palo Alto, California, 1983).
13. Prahm, L. P., Torp, U. & Stern, R. M. *Tellus* **28**, 355–372 (1976).
14. Wolff, G. T. *et al. Atmos. Envir.* **20**, 1229–1239 (1986).
15. Barrie, L. A. & Hoff, R. M. *Atmos. Envir.* **19**, 1995–2010 (1985).
16. Prospero, J. M., Savoie, D. L., Nees, R. T., Duce, R. A. & Merrill, J. *J. geophys. Res.* **90**, 10586–10596 (1985).
17. Savoie, D. L., Prospero, J. M. & Saltzman, E. S. Am. Chem. Soc., Division of Environmental Chemistry, Preprints of papers presented at the 194th National Meeting, New Orleans, Louisiana, 30 August–4 September, 1987; Vol. 27, no. 2, 324–327.
18. Lawson, D. R. & Winchester, J. W. *Science* **205**, 1267–1269 (1979).
19. Saltzman, E. S., Savoie, D. L., Prospero, J. M. & Zika, R. G. *J. atmos. Chem.* **4**, 227–240 (1986).
20. Andreae, M. O. *et al. J. atmos. Chem.* **6**, 149–173 (1988).
21. Berresheim, H., Andreae, M. O., Ayers, G. P. & Gillett, R. Am. Chem. Soc., Division of Environmental Chemistry, Preprints of papers presented at the 194th National Meeting, New Orleans, Louisiana, 30 August–4 September, 1987; Vol. 27, no. 2, 263–266.
22. Galloway, J. N., Likens, G. E., Keene, W. C. & Miller, J. M. *J. geophys. Res.* **87**, 8771–8786 (1982).
23. Nyberg, A. *Q. J. R. met. Soc.* **103**, 607–615 (1977).
24. Jickells, T., Knap, A., Church, T., Galloway, J. & Miller, J. *Nature* **296**, 55–57 (1982).
25. Whelpdale, D. M., Eliassen, A., Galloway, J. N., Dovland, H. & Miller, J. M. *Tellus* **40B**, 1–15 (1988).
26. Neftel, A., Beer, J., Oeschger, H., Zürcher, F. & Finkel, R. C. *Nature* **314**, 611–613 (1985).
27. Mayewski, P. A. *et al. Science* **232**, 975–977 (1986).
28. Barrie, L. A., Fisher, D. & Koerner, R. M. *Atmos. Envir.* **19**, 2055–2063 (1985).
29. Delmas, R. & Boutron, C. *J. geophys. Res.* **85**, 5645–5649 (1980).
30. Herron, M. M. *J. geophys. Res.* **87**, 3052–3060 (1982).
31. Cess, R. D. *J. atmos. Sci.* **33**, 1831–1843 (1976).
32. Ellis, J. S. thesis, Colorado State Univ. (1978).
33. Berlyand, T. G., Strokina, L. A. & Greshnikova, L. E. *Soviet Met. Hydrol.* no. 3, 9–15 (1980).
34. Ramanathan, V. *Science* **240**, 293–299 (1988).
35. Fishman, J., Ramanathan, V., Crutzen, P. J. & Liu, S. C. *Nature* **282**, 818–820 (1979).
36. Wigley, T. M. L. & Raper, S. C. B. *Nature* **330**, 127–131 (1987).
37. Hansen, J. *et al. in Climate Processes and Climate Sensitivity* Geophys. Monogr. **29**, Maurice Ewing Vol. 5 (eds Hansen, J. E. & Takahashi, T.) 130–163 (Am. geophys. Un. Washington DC, 1984).
38. Jones, P. D., Wigley, T. M. L. & Wright, P. B. *Nature* **322**, 430–434 (1986).
39. Jones, P. D. *et al. J. Clim. appl. Met.* **25**, 161–174 (1986).
40. Jones, P. D., Raper, S. C. B. & Wigley, T. M. L. *J. Clim. appl. Met.* **25**, 1213–1230 (1986).
41. Hansen, J. & Lebedeff, S. *J. geophys. Res.* **92**, 13345–13372 (1987).
42. Husar, R. B., Lodge, J. P. Jr & Moore, D. J. (eds) *Sulfur in the Atmosphere: Proc. int. Symp.* Dubrovnik, Yugoslavia, 7–14 September, 1977 (Pergamon, Oxford, 1978).
43. Ayers, G. P., Ivey, J. P. & Goodman, H. S. *J. atmos. Chem.* **4**, 173–185 (1986).
44. Cunningham, W. C. & Zoller, W. H. *J. Aerosol Sci.* **12**, 367–384 (1981).
45. Churning, T. M., Galloway, J. N., Jickells, T. D. & Knap, A. H. *J. geophys. Res.* **87**, 11013–11018 (1982).
46. Keene, W. C., Pszenny, A. A. P., Galloway, J. N. & Hawley, M. E. *J. geophys. Res.* **91**, 6647–6658 (1986).
47. Likens, G. E., Keene, W. C., Miller, J. M. & Galloway, J. N. *J. geophys. Res.* **92**, 13299–13314 (1987).

A human immunoglobulin gene reduces the incidence of lymphomas in c-Myc-bearing transgenic mice

Michel C. Nussenzweig*, Emmett V. Schmidt*, Albert C. Shaw*,
Eric Sinn*, Juanita Campos-Torres*, Bernard Mathey-Prevot†,
Paul K. Pattengale‡ & Philip Leder*

* Department of Genetics, Howard Hughes Medical Institute and Harvard Medical School, Boston, Massachusetts 02115, USA

† Pediatric/Oncology Department, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts 02115, USA

‡ Department of Pathology, Childrens Hospital of Los Angeles, Los Angeles, California 90027, USA

Transgenic mice carrying an immunoglobulin enhancer-driven c-myc oncogene develop rapid-onset pre-B cell lymphomas. The incidence of these malignancies is greatly reduced when an additional transgene encoding the membrane-bound form (but not the secreted form) of human Igμ is bred into the susceptible strain. This suppressive effect correlates with a subtle alteration in B-cell development induced by the immunoglobulin transgene.

It has long been suspected that a complex variety of genetically determined factors influences an individual's predisposition to the development of cancer¹. Although few malignancies follow strict Mendelian patterns of inheritance, several display clear familial associations². These associations are thought to be due in part to genes that segregate independently and contribute fractionally to risk. The nature of such genes has been a matter of considerable speculation, much of which deals with tumour-promoting and tumour-suppressing activities (for example, proto-oncogenes and anti-oncogenes^{3,4}). Some of these genes might however be indirectly involved in the development of malignancy, for example by modulating various developmental pathways or, more specifically, by altering the expression of the genes of the major histocompatibility complex⁵.

The availability of transgenic mice that develop predictable specific tumours enables genes to be identified that influence the occurrence of malignancy. In principle, the *in vivo* effect of any gene or transgene can be evaluated by combining the appropriate alleles using classical genetic techniques. To assess the effects of genetically induced developmental perturbations on transformation, we have crossed a transgenic line that develops pre-B cell lymphomas⁶ with strains of mice that carry human immunoglobulin transgenes. Our tester strain, a lymphoma-prone line, EnMyc, carries a human c-myc gene that has been activated by inserting a murine immunoglobulin heavy chain enhancer into its first intron⁶ (Fig. 1c). These transgenic mice are similar to those created by Adams *et al.*⁷ and develop clonal lymphomas which are predominantly pre-B cell in origin.

The pathway of B-cell development was perturbed in the tester strain by introducing a transgene encoding the membrane-bound form of the immunoglobulin heavy chain (mIgμ). Igμ is synthesized in two forms, membrane-bound and secreted^{8,9}. Only the membrane-bound form is active in suppressing further rearrangements of immunoglobulin heavy-chain genes¹⁰, a phenomenon known as allelic exclusion. B cells that express the rearranged mIgμ transgene bypass an early stage in their development because they produce biologically active heavy chains without undergoing heavy-chain gene rearrangements. Indeed, the rearrangement of endogenous heavy-chain genes is suppressed by the mIgμ transgene^{11,12}. We show here that the mIgμ transgene also suppresses the development of early onset pre-B lymphomas in mice that carry the human c-myc transgene. The same Igμ transgene modified to produce only the secreted form of the heavy chain (sIgμ) has no effect on either heavy-chain gene rearrangements^{13,14} or the incidence of lymphoma.

Membrane IgM suppresses lymphoma

To produce transgenic mice that carry both the EnMyc transgene and either the membrane-bound or the secreted version of the human Igμ transgene, heterozygous carrier mice were mated^{6,10,14} (Fig. 1c). Dual carrier progeny (EnMyc/mIgμ and EnMyc/sIgμ), identified by DNA tail blots, were produced almost at the expected Mendelian frequency of 25%. Mice were inspected on a weekly basis for early signs of lymphoma, usually manifested by lymph node enlargement. These experiments were carried out on a largely CD-1 outbred genetic background using littermates as controls. As previously documented, EnMyc transgenic mice develop rapid onset pre-B cell lymphoma with high penetrance⁶ (Fig. 1); 50% of the animals carrying this gene develop tumours by 117 days of age ($T_{50} = 117$ days) (Fig. 1a). During this same period, none of their littermates carrying both the membrane-bound form of human Igμ and EnMyc transgenes developed tumours. After one year, more than 90% of the EnMyc mice had developed tumours as compared with only 40% of the EnMyc/mIgμ double carriers (Fig. 1a). Thus, the membrane-bound form of the human Igμ transgene is associated with a delay and reduction in the incidence of EnMyc-induced lymphomas. In contrast, the incidence of lymphomas in carriers of both the secreted Igμ and EnMyc transgenes is virtually identical to that of littermates which carry the EnMyc transgene alone (Fig. 1b). As protection was also seen with another independently derived strain carrying the membrane-bound transgene (data not shown), it is unlikely that either the site of integration of the heavy chain gene or disruption of closely associated genes is responsible for the differences between EnMyc/sIgμ and EnMyc/mIgμ mice.

mIgμ transgene alters tumour phenotype

Although the incidence of tumours is reduced in EnMyc/mIgμ carrier mice, tumours do arise after a long latency period. These tumours are phenotypically different from those in control EnMyc and EnMyc/sIgμ mice. By staining primary tumour cells with a panel of antibodies specific for cell-surface antigens, tumours from each set of animals were assigned to one of five groups (Fig. 2). Eighty-three per cent of 50 consecutive tumours from EnMyc animals were of the pre-B cell type (Fig. 2a). The remainder were either mature B cells (6%), T cells (2%) or null cell lymphomas (8%). The majority of the tumours that arose in the control dual carriers of EnMyc/sIgμ ($n = 17$) were also pre-B cell lymphomas (88%; Fig. 2b). In contrast, the late-arising tumours from EnMyc/mIgμ mice ($n = 21$) were evenly

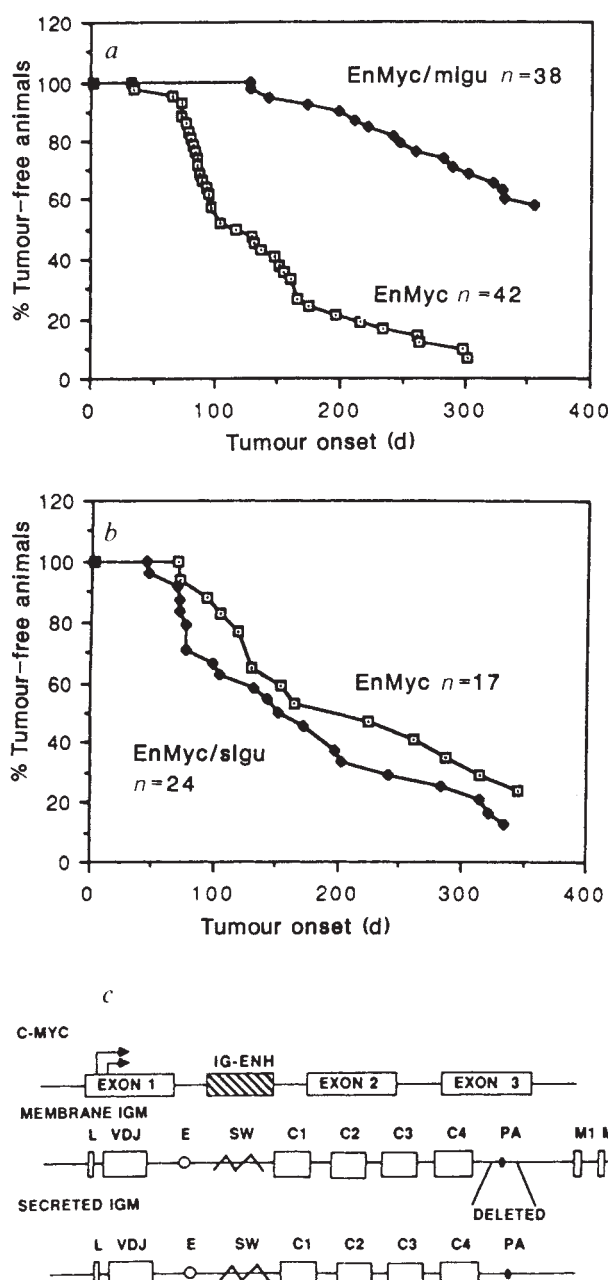


Fig. 1 Tumour occurrence as a function of age in transgenic mice. Tumour onset in days is plotted against the percentage of animals that remain tumour-free. *a*, Kinetics of tumour occurrence in EnMyc animals and their EnMyc/mIgμ littermates. *b*, Kinetics of tumour occurrence in EnMyc animals and their EnMyc/sIgμ littermates. *c*, Maps of *c-myc* and immunoglobulin transgenes. There were no lymphomas in the immunoglobulin transgenics or wild-type mice during these experiments. The human *c-myc* gene has a 1-kb *Xba*I fragment of the mouse immunoglobulin heavy chain enhancer inserted in the *Xba* site in the first intron⁶. The gene encoding only the membrane-bound version of the human Igμ chain was constructed by deleting the 5' polyadenylation site that is required for the formation of secreted Igμ (sIgμ) (refs 10 and 36). The gene encoding the secreted version of the Igμ chain was made by deleting the membrane exons from the same human μ gene that was used to make the mIgμ (ref. 14). Symbols: Ig-enh, 1 kilobase-long *Xba*I fragment of the mouse immunoglobulin heavy-chain enhancer⁶; L, leader exon; VDJ, recombined variable region; E, human heavy-chain enhancer; sw, immunoglobulin switch region; C1-C4, immunoglobulin constant region exons; pA, polyadenylation/cleavage site; M1 and M2, exons encoding peptides that anchor the Igμ product to the cell surface membrane.

Table 1 Histopathology of malignant lymphomas occurring in mice bearing the EnMyc, EnMyc/mIgμ, and EnMyc/sIgμ transgenes

Genotype	Sample size	AML	FCCL	IL	LL
EnMyc	9	1	1	—	7
EnMyc/mIgμ	9	2	3	2	2
EnMyc/sIgμ	14	3	1	—	11

A randomly selected set of 32 tumours was examined and categorized as acute myelogenous leukaemia (AML), follicular centre cell lymphoma (FCCL), immunoblastic lymphoma (IL) and lymphoblastic lymphoma (LL), according to criteria previously described¹⁵.

distributed among pro-B cells (29%), pre-B cells (24%) and mature B cells (24%) with a small proportion of T cell (5%), or null cell (14%) lymphomas (Fig. 2c).

Lymphomas derived from the EnMyc/mIgμ mice also differed histologically from lymphomas arising in mice that carry the EnMyc transgene alone. Whereas the majority of the malignancies in EnMyc mice are lymphoblastic lymphomas⁶ (Table 1), those occurring in EnMyc/mIgμ mice are evenly distributed among several histological types, including acute myelogenous leukaemia, follicular centre cell lymphoma, immunoblastic lymphoma, and lymphoblastic lymphoma¹⁵. Thus, the transgene encoding the membrane-bound form of Igμ not only delays the onset of malignancy and reduces its incidence, but also influences the phenotype of the tumours that do occur. This change in tumour phenotype results largely from a specific reduction of pre-B lymphoblastic malignancies because (with the exception of pro-B cell lymphomas) other classes of lymphoma continue to occur at similar absolute frequencies (Fig. 2).

Both transgenes expressed in target cells

As the heavy-chain enhancer is a key regulatory element in both the Igμ and EnMyc transgenes, the reduced incidence of tumours in dual carriers could be due to competition for enhancer factor(s) between the two transgenes. If this were so, one would have expected tumour suppression in the sIgμ (10–20 transgene copies) as well as the mIgμ (2–4 copies) dual carrier mice. This was clearly not the case (Fig. 1), so competition for enhancer factor(s) between transgenes is unlikely to explain tumour suppression. Using RNase protection, the steady-state level of human *c-myc* mRNAs in tumours from the three types of transgenic mice was assayed and found to be essentially equivalent (Fig. 3a). Note that a mouse β₂-microglobulin probe served as a control for the amount of mRNA in each sample. All 30 tumours expressed similar quantities of transgenic human *c-myc*, regardless of the type of lymphoma, and regardless of whether the tumour arose in mice carrying only EnMyc or a combination of EnMyc with one of the Igμ heavy-chain transgenes (see examples in Fig. 3a, top panel). In addition, an Igμ probe that distinguishes between the products of the mIgμ and sIgμ transgenes was used to assay human Igμ expression (see hu-CH4, Fig. 3). As expected, all the lymphomas in dual carrier mice expressed their human Igμ transgenes (Fig. 3a, bottom panel).

To determine whether EnMyc expression was reduced in B-cell precursors, we examined mRNA levels in bone marrow samples enriched in pre-B cells (Fig. 3b). The number of pre-B cells in the samples was determined by flow cytometry and found to be highly variable. In 16 independent samples, however, the level of transgenic *c-myc* mRNA was directly related to the number of pre-B cells and was not affected by the presence or absence of the mIgμ (see examples in Fig. 3b). As we would expect if enhancer factors are not limiting, human Igμ mRNA expression paralleled transgenic *c-myc* mRNA expression (compare top and bottom panels in Fig. 3b). It is thus unlikely that protection from lymphoma depends on a reduction in *c-myc* mRNA in pre-B cells.

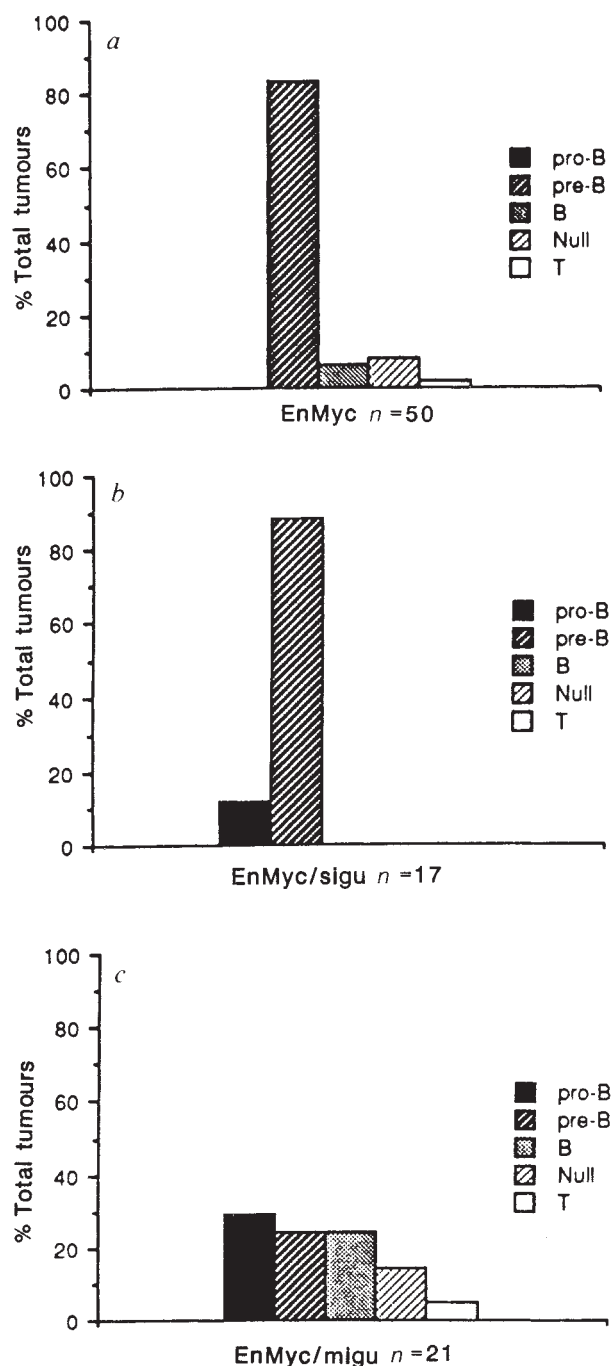


Fig. 2 Cell-surface antigens of transgenic tumours. All tumours arising within a period of 12 months were analysed. The percentage of the total number of tumours in each of five categories is plotted for EnMyc, EnMyc/sigm, and EnMyc/migm animals. Tumour cells were prepared and stained for fluorescence analysis on an Ortho Cytofluorograf II as previously described¹⁴. The cells were stained with either fluorescein-labelled primary antibodies to mouse μ (ref. 37), mouse κ (Southern Biotechnology), Thy1.2 (ref. 38), Ly-2 (ref. 38) or with culture supernatants from hybridomas producing anti-L3T4 (ref. 39), anti-Mac-1 (ref. 40), anti-Ly-5 (B220; ref. 41), anti-ThB (ref. 38), or anti-human μ (ref. 42), and developed with either fluorescein-labelled goat anti-mouse IgG1, or goat anti-rat IgG (Southern Biotechnology). B-cell tumours were positive for surface immunoglobulin and Ly-5 or ThB. Pre-B lymphomas were positive for Ly-5 or ThB and negative for surface κ and Ig μ . T-cell lymphomas were positive for Thy1.2 and either L3T4 or Ly-2. Myelomonocytic tumours were positive for Mac-1 and negative for Ly-5, ThB and Thy1.2. Null tumours were those that did not express either Ly-5, ThB, Thy1.2 or Mac-1. Pro-B cell tumours express Mac-1 and Ly-5, grow in Dexter⁴³ culture conditions and develop into B cells in Whittlock-Witte cultures⁴⁴.

Table 2 A-MuLV transformation of bone marrow cells in soft agar

Experiment number	Genotype	Mouse number	Colonies
1	mIg μ	1,571	6; 13
	mIg μ	1,573	15; 18
	Wt	1,576	60; 61
	Wt	1,577	86; 110
2	mIg μ	1,624	0
	Wt	1,625	64
3	mIg μ	1,709	40; 39
	mIg μ	1,713	21; 23
	Wt	1,710	230; 250
	Wt	1,711	157; 175

Bone-marrow cells from mIg μ transgenic mice or their wild-type (Wt) littermates were infected in duplicate with Abelson virus and plated in soft agar as previously described¹⁷. Colonies were counted after 2 weeks. Independent viral stocks with different titres were used in each experiment.

EnMyc is expressed in pre-B cells

Earlier studies have shown that EnMyc mRNA is abundant in bone marrow but not spleen cells⁶, suggesting that this transgene is predominantly expressed in the earlier stages of B-cell development. To document this directly, steady-state levels of human c-myc mRNA were measured in purified preparations of mature B and pre-B cells. As shown by RNase protection analyses (Fig. 3c), isolated pre-B cells from bone marrow express higher levels of transgenic c-myc than isolated mature B cells. This high expression of the transgene in the pre-B cell stage of development could account for the predominant occurrence of pre-B cell lymphomas in EnMyc mice. Moreover, it is during this pre-B cell stage that the membrane-bound heavy chain is active in excluding the rearrangement of endogenous mouse heavy-chain genes¹⁶.

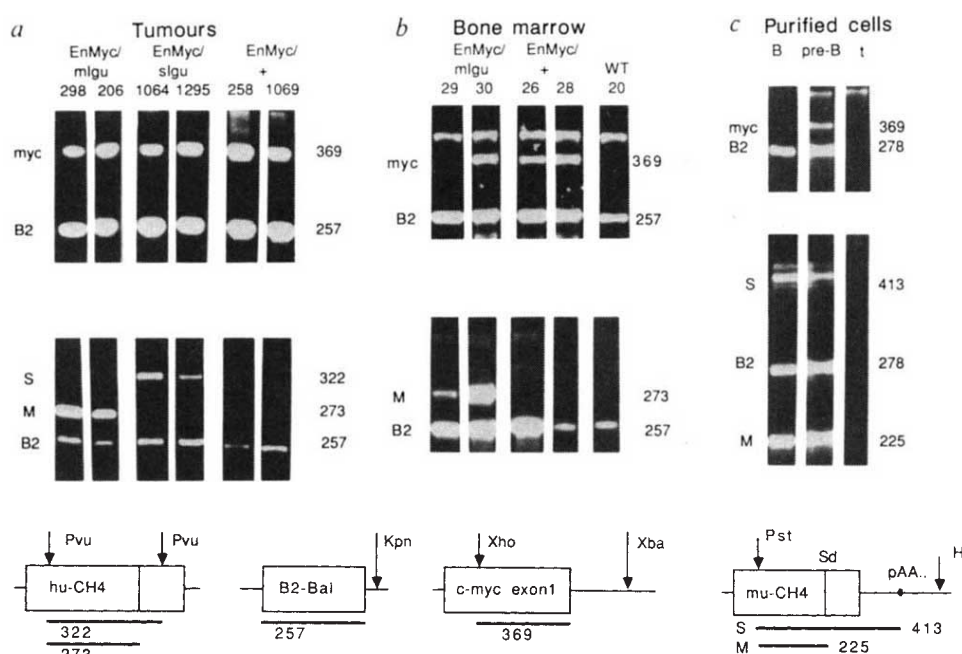
Immune surveillance

The introduction of human cell-surface mIg μ chains into a mouse raises the issue of a direct role for the immune system in tumour suppression. As transgenic mice carrying the human membrane μ gene are tolerant to human μ (it is expressed by day 9 of embryonic development), and human μ appears on the surface of most of the B cells in mice that carry this transgene¹⁰, immune rejection is very unlikely to account for the reduced and delayed tumour incidence we observe. Further, mIg μ transgenic mice have normal numbers of B cells and have shown no clinical or pathological evidence of autoimmune disease in over two years of observation. We confirmed that the immune system *per se* does not have a critical role in tumour suppression by reproducing our *in vivo* observations *in vitro* in the absence of an intact immune system, using the Abelson murine leukaemia virus (A-MuLV) to infect bone marrow cells (Table 2).

Sensitivity to A-MuLV transformation

One explanation for the tumour-suppressing effect of membrane-bound μ is that expression of the transgene physiologically alters the pre-B cell. Although fluorescence-activated cell sorter analyses do not detect any consistent differences in the numbers of pre-B cells in 40 independent bone marrow samples from immunoglobulin-carrying transgenic mice and their matched wild-type littermates (not shown), the number of pre-B cells in normal bone marrow is small and the sensitivity of the assay is such that genetic variability and immunological status may account for the variation in pre-B cell numbers. It is also possible that pre-B cells from mIg μ transgenics differ from normal pre-B cells in a way not measured by conventional surface antigen analysis. An alternative way to assess the trans-

Fig. 3 Expression of *c-myc* and human $\text{Ig}\mu$ in tumours and bone marrow cells. Expression of mRNA was assayed by RNase protection⁴⁵. Tumour RNA (10 μg) was annealed with the appropriate anti-sense RNA probe at 53 °C overnight. Single-stranded RNA was digested with a mixture of ribonuclease A and ribonuclease T1. Protected fragments were visualized by autoradiography of 5% acrylamide/7M urea sequencing gels. A mouse β_2 -microglobulin probe was included in all experiments to control for the level of mRNA in the samples⁴⁶. Tumour number and genotypes are indicated above each lane. Transfer RNA controls were run for each analysis, but are not shown except in panel c. Undigested probe is visible above the 369-base pair *c-myc* band in some lanes (see tRNA in c). *a*, Human *c-myc*, human $\text{Ig}\mu$ and mouse β_2 microglobulin expression in tumours. *b*, Human *c-myc*, human $\text{Ig}\mu$, and mouse β_2 -microglobulin expression in bone marrow cells. To



determine the number of pre-B cells, an aliquot of each bone marrow sample was stained with either anti-Ly-5 (B220, ref. 20) and developed with fluorescein-labelled goat anti-rat, or stained with fluorescein labelled anti-mouse κ . The number of Ly-5-positive kappa-negative pre-B cells was; 5% in number 20, 35% in number 26, 30% in number 28, 2.4% in number 29, and 31% in number 30. *c*, Human and mouse *c-myc* β_2 microglobulin expression in mature B cells and pre-B cells derived from EnMyc transgenic mice. Upper panel shows the human *c-myc* transgene; lower panel, endogenous mouse $\text{Ig}\mu$. The β_2 microglobulin probe used in the latter experiment protects a slightly longer fragment (278 nucleotides) than the probe used in experiments shown in Fig. 3*a* and *b* (ref. 10). Mature B cells were prepared by 'panning' for mouse IgM -positive spleen cells⁴⁷, and pre-B cells were prepared by first depleting bone marrow cells of IgM positive cells, and then panning for Ly-5-positive cells on dishes coated with anti-rat immunoglobulin. The RNase protection probes are shown in the diagrams (bottom), with the expected protected fragments indicated in nucleotides. Symbols: β_2 , β_2 -microglobulin; S, secreted $\text{Ig}\mu$; M, membrane bound $\text{Ig}\mu$; *c-myc*, human *c-myc*; h μ -CH4, human $\text{Ig}\mu$ fourth constant region exon; M μ -CH4, mouse $\text{Ig}\mu$ fourth constant region exon; Sd, splice donor; Pvu, *Pvu*II site; Kpn, *Kpn*I site. The size of protected fragments is given in bases. Undigested full-length probes are seen in upper panels *a*, *b* and *c* as unlabelled bands.

genic pre-B cell population is to assay susceptibility to Abelson virus transformation¹⁷. The number of soft agar colonies obtained post-infection is an indirect measure of the number of pre-B cells that are susceptible to transformation by the oncogene *v-abl*. In five sets of matched littermates, we found that mice carrying the m $\text{Ig}\mu$ transgene produce 70–80% fewer A-MuLV bone marrow transformants than wild-type (Table 2). This result indicates that pre-B cells from m $\text{Ig}\mu$ transgenic mice are less sensitive to the transforming properties of *v-abl* *in vitro*, just as they are less sensitive to the transforming properties of transgenic EnMyc *in vivo*.

Tumour suppressing mechanism

We have already noted that the major target of the EnMyc gene is the pre-B cell, a category of cells that is committed to B-cell development and expresses a defined set of surface antigens. Pre-B cells have undergone at least the first of two rearrangements necessary to produce an active immunoglobulin heavy-chain gene, but have not formed an active light-chain gene and hence do not produce surface immunoglobulin¹⁶. It is quite probable that there are developmental substages within this category, at the least a stage between the beginning of immunoglobulin heavy-chain gene rearrangement and the appearance of membrane-bound heavy-chain protein within the pre-B cell's endoplasmic reticulum. The intact membrane-bound heavy chain is a part of the biological effector system that mediates allelic exclusion^{10–12,18}. Its appearance most probably heralds an alteration in the physiology of the pre-B cell, albeit one that cannot be detected by the surface antigens currently used to classify pre-B cells. The active m $\text{Ig}\mu$ transgene would be expected to accelerate the passage of the pre-B cell through this developmental pathway¹⁹. Indeed, although all immunoglobulin heavy-chain alleles are rearranged in EnMyc lym-

phomas⁶, 13 out of 24 alleles from m $\text{Ig}\mu$ /EnMyc tumours were in the germline configuration (not shown).

The foregoing leads us to suggest that the m $\text{Ig}\mu$ transgene exerts its effect on oncogenesis by accelerating the development of the pre-B cell, thus reducing the size of the population of cells most susceptible to transformation by EnMyc. The observation that the control slg μ transgene has no effect on allelic exclusion and hence no effect on pre-B cell development or oncogenesis is consistent with this idea. In contrast, the Epstein-Barr virus, which is associated with the endemic Burkitt lymphoma in young African children²⁰, could exert its postulated tumour-promoting effect by enlarging the specific population of cells at risk for the Burkitt translocation, in effect causing a pause in the B-cell developmental pathway.

Tumour suppressors, recessive or anti-oncogenes

In our transgenic paradigm the m $\text{Ig}\mu$ transgene acts in a dominant fashion in that a single chromosome carrying the transgene locus protects against the development of lymphoma. Nonetheless, the m $\text{Ig}\mu$ transgene can also be thought of as a recessive anti-oncogene because its elimination from the genome restores the tumorigenic phenotype. This recessive genetic behaviour resembles what has been described for retinoblastoma^{21–23}, Wilms' tumour^{24–27}, acoustic neuroma²⁸, colorectal carcinoma²⁹, and the lethal(2) giant larvae tumour suppressor gene in *Drosophila*³⁰. The retinoblastoma and lethal(2) giant larvae genes have been identified and sequenced^{31–34} and although the putative coding sequences of these genes do not resemble known proteins, it has been postulated that these and other anti-oncogenes are involved in developmental regulation^{1,31,35}. Our observations suggest that the protective effect of the m $\text{Ig}\mu$ transgene is due to a physiological alteration of pre-B cells, the population of cells at highest risk for transformation.

This example implies that a diverse group of genes, not usually considered to be oncogenes or anti-oncogenes but known to control development, can influence both the incidence and rate of onset of malignancy. Though loss of gene function may directly cause malignant transformation in other systems, an alternative possibility is that an allele inducing a developmental pause or acceleration may be a more general,

albeit indirect, mechanism that influences susceptibility to malignancy.

We thank Dr Naomi Rosenberg for Abelson virus stock, Ms Racheal Wallace for supervision of the mouse experiments and Drs Daniel Sussman and Akira Shimizu for discussion. This work was supported in part by a grant from E.I. Du Pont de Nemours and Co. Inc.

Received 15 August; accepted 25 October 1988.

- Hansen, M. F. & Cavenee, W. K. *Cancer Res.* **47**, 5518-5527 (1987).
- Muller, H. & Weber, W. (eds) *Familial Cancer* (Karger, Basle, 1985).
- Bishop, J. M. *Science* **235**, 305-311 (1987).
- Knudson, A. G. *A. Rev. Genet.* **20**, 231-251 (1986).
- Tanaka, K., Isselbacher, K. J., Khoury, G. & Jay, G. *Science* **228**, 26-30 (1985).
- Schmidt, E. V., Pattengale, P., Weir, L. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6047-6051 (1988).
- Adams, J. M. *et al. Nature* **318**, 533-538 (1985).
- Rogers, J. *et al. Cell* **20**, 303-312 (1980).
- Alt, F. W. *et al. Cell* **20**, 293-302 (1980).
- Nussenzweig, M. C. *et al. Science* **236**, 816-819 (1987).
- Weaver, D., Costantini, F., Imanishi-Kari, T. & Baltimore, D. *Cell* **42**, 117-127 (1985).
- Rusconi, S. & Kohler, G. *Nature* **314**, 330-334 (1985).
- Storb, U. *et al. J. exp. Med.* **164**, 627-641 (1986).
- Nussenzweig, M. C., Shaw, A. C., Sinn, E., Campos-Torres, J. & Leder, P. *J. exp. Med.* **167**, 1969-1974 (1988).
- Pattengale, P. P. & Taylor, C. R. *Am. J. Pathol.* **113**, 237-265 (1983).
- Coffman, R. L. & Weissman, I. L. *J. molec. Immun.* **1**, 31-38 (1983).
- Rosenberg, N. & Baltimore, D. *J. exp. Med.* **143**, 1453-1463 (1976).
- Alt, F. W. *et al. EMBO J.* **3**, 1209-1219 (1984).
- Reth, M., Petrac, E., Wiese, P., Lobel, L. & Alt, F. W. *EMBO J.* **6**, 3299-3305 (1987).
- de The, G. *et al. Nature* **274**, 756-761 (1978).
- Cavenee, W. K. *et al. Nature* **305**, 779-784 (1983).
- Godbout, R., Dryjia, T. P., Squire, J., Gallie, B. & Phillips, R. A. *Nature* **304**, 451-453 (1983).

- Sparkes, R. S. *et al. Science* **219**, 971-973 (1983).
- Koufos, A. *et al. Nature* **309**, 170-172 (1984).
- Orkin, S. H., Goldman, D. S. & Sallan, S. E. *Nature* **309**, 172-174 (1984).
- Reeve, A. E. *et al. Nature* **309**, 174-176 (1984).
- Fearon, E. R., Vogelstein, B. & Feinberg, A. P. *Nature* **309**, 176-178 (1984).
- Seizenger, B. R., Martuza, R. L., & Gusella, J. F. *Nature* **322**, 644-647 (1986).
- Solomon, E. *et al. Nature* **328**, 616-619 (1987).
- Gateff, E. *Science* **200**, 1448-1459 (1978).
- Friend, S. H. *et al. Nature* **323**, 643-646 (1986).
- Lee, W. H. *et al. Science* **235**, 1394-1399 (1987).
- Fung, Y. K. T. *et al. Science* **236**, 1657-1661 (1987).
- Jacob, L., Oppen, M., Metzroth, B., Phannavong, B. & Mechler, B. M. *Cell* **50**, 215-225 (1987).
- Hansen, M. F. & Cavenee, W. K. *Cell* **53**, 172-173 (1987).
- Danner, D. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8658-8662 (1985).
- Kung, J., Sharrow, S. O., Sieckman, D. G., Liberman, R. & Paul, W. E. *J. Immun.* **127**, 873-876 (1981).
- Ledbetter, J. A. & Hertzberg, L. A. *Immunol. Rev.* **47**, 63-90 (1979).
- Dialynas, D. P. *et al. Immunol. Rev.* **74**, 29-56 (1983).
- Springer, T., Galfre, G., Sechere, D. S. & Milstein, C. *Eur. J. Immun.* **9**, 302-306 (1979).
- Coffman, R. A. *Immunol. Rev.* **69**, 5-23 (1982).
- Maruyama, S., Kubagawa, H. & Cooper, M. D. *J. Immun.* **135**, 192-199 (1985).
- Dexter, T. M., Allen, T. D. & Lajtha, L. G. *J. cell. Physiol.* **91**, 335-344 (1977).
- Whitlock, C. A. & Witte, O. N. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3608-3612 (1982).
- Melton, D. A. *et al. Nucleic Acids Res.* **12**, 7035-7055 (1984).
- Parnes, J. R. & Seidman, J. G. *Cell* **29**, 661-669 (1982).
- Wysocki, L. J. & Sato, V. L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2844-2848 (1978).

LETTERS TO NATURE

Delayed hard X-rays from Cygnus X-1

Signori Miyamoto*, Shunji Kitamoto*,
Kazuhisa Mitsuda† & Tadayasu Dotani†

* Department of Physics, Faculty of Science, Osaka University, 1-1, Machikaneyama, Toyonaka, Osaka 560, Japan

† Institute of Space and Astronautical Science, 3-1-1, Yoshinodai, Sagami-hara, Kanagawa 229, Japan

Cygnus X-1, a black hole candidate of long standing, emits X-rays whose flux varies on many timescales^{1,2}. The X-rays are commonly thought to be due to inverse Compton scattering of low-energy photons by a cloud of hot plasma, with temperature $\sim 10^8$ - 10^9 K, near the black hole^{3,4}. According to this model, variations in the flux of high-energy X-rays should occur somewhat later than similar variations in lower-energy X-rays. Using the large-area proportional counters (LAC) from the Ginga satellite, we have observed Cygnus X-1 in its low state, and complex cross-spectrum analysis of the data reveals that the delay of the high-energy X-rays with respect to the low increases from approximately 2 ms to several seconds for periods from 0.1 to 300 s. This is inconsistent with the inverse Compton scattering model of Cygnus X-1.

It is believed that in Cyg X-1 matter from the donating O-type supergiant forms an accretion disk around the black hole and that the gravitational energy of the matter in the disk is transformed into thermal energy¹. It has been suggested that the high-energy X-rays from Cyg X-1 are the result of high-energy-gain inverse Compton scattering of low-energy photons by a hot plasma cloud situated on the central part of the disk^{3,4}. The temperature of the hot plasma, T_e , and the Thomson scattering depth are estimated to be $\sim 3 \times 10^8$ K and 5 respectively⁴.

The inverse Compton scattering model predicts that, if there is a time variation in the initial low-energy photons, the corresponding time variation of the high-energy Compton X-rays will be delayed relative to that of the low-energy X-rays. There have been several attempts⁵⁻¹⁰ to look for time delays using cross-correlation methods. Page *et al.*⁹ found a time lag of 7.5 ms

between X-rays of energy 3-40 keV and those of energy 1-3 keV during the high state of Cyg X-1, and Page¹⁰ reported a lag of about 6 ms between 5-14-keV and 2-5-keV X-rays in the low state.

Recently, van der Klis *et al.*¹¹ calculated complex cross-spectra of Cyg X-2 and GX 5-1, and found hard X-ray time-lags of ~ 2 and 0.4 ms, respectively, at the frequencies of the quasisperiodic oscillations (QPO). A soft time-lag at QPO frequencies was observed in the Rapid Burster¹². These results inspired us to examine the time behaviour of Cyg X-1 using the same method.

Observations were made from 5 to 8 August 1987 with the Ginga large-area proportional counters (LAC)¹³. On 5 August, the observations were made in the PC mode and the observed energy ranges (and time resolutions) were 1.2-5.7 keV (1 ms), 5.7-24.4 keV (2 ms), 1.2-15.8 keV (1 ms) and 15.8-24.4 keV (2 ms); the effective area for each energy range was about 2,000 cm². From 6 to 8 August, the observations were made in the MPC-3 mode (12 channels in the 1.2-37-keV band with a time resolution of 8 ms), with an effective area of $\sim 4,000$ cm².

This report is based on the observations of 5 and 6 August, for which orbital phases were 0.6 and 0.8, respectively. During these observations, Cyg X-1 was in its low state. To measure the delay time between X-rays of various energy in various time periods, we used the complex cross-spectrum analysis technique described by van der Klis *et al.*¹¹.

We have computed the complex cross-spectra and studied the phase and time delay between two energy bands, as a function of the Fourier periods from 4 ms to 400 s, using two separate sets of data. The first set consists of 552 16-s segments. Each segment is further binned into smaller elements with a time resolution of 2 ms. The second set consists of fourteen 400-s segments binned with a time resolution of 0.8 s. The complex Fourier transforms for each of the two energy bands in each data set were computed using the fast Fourier transform technique. These complex Fourier amplitudes were averaged over the 552 (or 14) segments. The relative phase (the phase lag) and the time delay (the relative phase multiplied by the period) were calculated from these averaged complex Fourier amplitudes. Finally, we averaged again the phase lags and time delays of

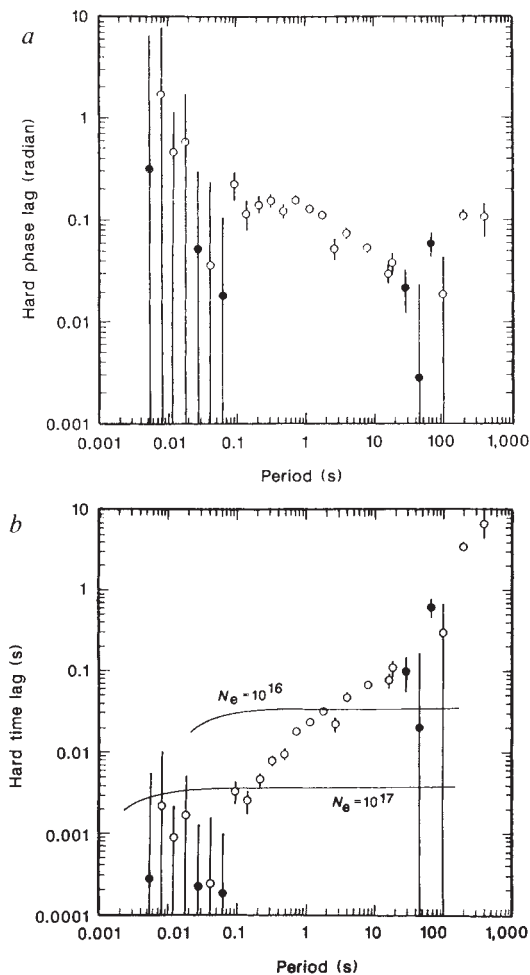


Fig. 1 The phase lag (a) and the time lag (b) between the hard X-rays (15.8–24.4 keV) and the soft X-rays (1.2–5.7 keV) (5 August data). Open circles denote positive values and filled circles correspond to negative values (soft-X-ray lag). Similar results are obtained between other X-ray energy ranges and from the data of other observations. Also shown in b are the simulated time delays between X-rays of energies 3.5 keV and 20 keV, where we have assumed that $\tau_{\text{es}} = 5$, $kT = 27$ keV, $N_e = 10^{16}$ and 10^{17} cm^{-3} , and the initial photon energy is 100 eV. Except for N_e , the parameters are those proposed by Sunyaev and Trumper⁴.

each Fourier component within a certain period range, with a width selected to increase the period progressively by a factor of 1.5, for plotting at equal intervals on a log scale.

The results for the PC mode data are shown in Fig. 1. In the region of overlap (periods of 1.6–16 s) between the two data sets, the results are consistent with each other, and for this region we have plotted in Fig. 1 only the results obtained from the 16-s segments. The errors shown are 1σ , estimated by the method described by Bendat and Piersol¹⁴. The dead-time-induced channel cross-talk mentioned by van der Klis *et al.*¹¹ is negligible in our case.

The results of this analysis can be summarized as follows. (1) Almost all of the hard X-ray phase-lags between 15.8–24.4 keV and 1.2–5.7 keV are in the range 0.01–1.0 radian. (2) The hard X-ray delay time between 15.8–24.4-keV and 1.2–5.7-keV X-rays is ~ 2 ms for a period of ~ 0.1 s, and the delay time increases almost linearly up to several seconds for a period of ~ 300 s. (3) For periods > 20 s, soft X-ray lags appear frequently. (4) The range of periods for which soft X-ray lag occurs seems to be the same even for different pairs of the two X-ray energy bands. However, this range changes with time: the range of the soft X-ray lag for 5 August is different from that for 6 August. (5) The lag occurs over the complete range of energy encom-

passed by the observations. As the energy separation of the two bands increases the hard X-ray lag increases. On average, the phase-lags between the 5.7–24.4-keV and 1.2–5.7-keV X-rays are smaller than those between the 15.8–24.4-keV and 1.2–5.7-keV X-rays by a factor of ~ 1.4 for both the 5 and 6 August data.

We now examine the consequences of these results for the inverse Compton scattering model of high-energy X-ray production in Cyg X-1. We use the formulae derived by Sunyaev and Titarchuk¹⁵ and Payne¹⁶ to estimate the period-dependence of the hard X-ray time-lag.

Assuming a spherical hot plasma of a temperature T , Payne¹⁶ derived the time t_{max} taken for an instantaneous input of soft photons of energy E_0 to reach maximum intensity at a given energy E :

$$t_{\text{max}} \approx \frac{1}{2} \left(\frac{9}{4} + \left(\frac{4\alpha}{y} \right) \right)^{-1/2} \frac{\ln(E/E_0)}{(kT/m_e c^2) N_e \sigma_T c}$$

where α is a geometrical factor of the order of a few, $y = (4kT/m_e c^2) \tau_{\text{es}}^2$ (τ_{es} being the Thomson scattering depth), m_e is the mass of the electron, N_e is the electron number density and σ_T is the Thomson scattering cross-section.

The Compton-scattered X-ray flux decreases exponentially with an energy-independent decay time constant of $\sim 3R\tau_{\text{es}}/(\pi^2 c)$, where R is the radius of the hot cloud¹⁵. As this decay time constant is independent of the X-ray energy, the time delay is determined mainly by t_{max} . Thus it is expected that the time delay between X-rays of different energy will be similar for all periods.

We confirmed this using Payne's formula¹⁶, from which we obtained the X-ray light curve that emerges from the hot plasma cloud for the case of an instantaneous delta-function, monochromatic photon source. Through standard Fourier transform techniques we calculated the complex cross-spectra between the two Compton-scattered light curves of different energies, and obtained the period-dependence of the time delay. As expected, the time delay is almost independent of the period (Fig. 1b). Thus the substantial period-dependence of the time lag which is observed cannot be explained by inverse Compton scattering.

One might suggest that the delay time of ~ 2 ms, observed for a period of < 0.1 s, is due to inverse Compton scattering, and that the strong period-dependence arises before this scattering. This cannot be the case, however, because the energy gained by the photons through scattering is very large compared with their initial energy^{3,4}, and this energy gain will fluctuate; any time-delay relations which exist amongst the initial photons would be washed out by the time the photons escape from the hot cloud. An initial time-delay relation could be preserved only if the Thomson scattering depth of the hot cloud is small and if the net energy change through scattering is small relative to the initial photon energy, neither of which is the case for the inverse Compton scattering model^{3,4}.

If high-energy X-rays in Cyg X-1 are produced mainly by high-energy-gain inverse Compton scattering, some modulation mechanism must be invoked to produce the strongly period-dependent time lag after scattering; there is no obvious source of such a mechanism. It is more likely that the phenomenon originates in the high-energy X-ray production mechanism itself. For example, vibration of the hot clouds at various frequencies could produce the observed time-lag/period relation in the emitted X-rays if the vibrations produce almost the same phase lags in the hard X-rays at the corresponding frequencies. Such a cloud-oscillation model has been discussed by Stollman *et al.*¹⁷, and is clearly different from the 'static' inverse Compton scattering models^{3,4}.

Alternatively the period-dependent delays could be due to the X-ray emission mechanism in the accretion disk. Suppose a clump of matter in the disk (which can emit more radiation because of its higher density) drifts from the outer (soft-X-ray-emitting) region to the inner (hard-X-ray-emitting) region. The drift time from the former to the latter region would then be

the cause of the time delay. The observed time-delay/period relation would then be explained if the drift time increases with the size of the clump.

Another possibility is that the time lag corresponds to the propagation time of a perturbing wave travelling from a low-energy- to a high-energy-X-ray-emitting region in the accretion disk with a group velocity $d\omega/dk$. Then the time delay corresponds to the wave travel-time. In this case it follows from the results in Fig. 1 that $d\omega/dk = a\omega^\gamma$, where ω is the angular frequency, k is the wave number, a is a constant and the value of γ is estimated at ~ 0.5 – 0.9 .

The soft-X-ray lag could occur by cooling of the X-ray-emitting plasma, or it could arise from some reverse perturbing wave travelling from the inner part of the disk to the outer part. The fact that the soft lag occurs in particular frequency regions would then imply that the reverse wave occurs in those regions.

Van der Klis *et al.*¹¹ observed soft-X-ray lags in Cyg X-2 and GX 5-1 at frequencies of less than ~ 7 Hz, in addition to a hard-X-ray lag at QPO frequencies. Although the absolute values of these time lags are of the same order of magnitude as the hard X-ray time lags of Cyg X-1 at the corresponding frequencies, they have opposite signs. Those authors proposed a model in which the hard-X-ray lag at QPO frequencies is due to Compton scattering by a hot electron cloud and the soft lag at low frequencies is due to the spectral softening of X-ray shots; the softening exceeds the Compton-induced hardening at these low frequencies¹¹. Stella *et al.*¹² observed soft lags of ~ 8 ms in the ~ 4 -Hz QPO during persistent emission intervals in the Rapid Burster, which may be caused by Compton scattering by a cold electron cloud. Mitsuda *et al.*¹⁸ observed a 70-ms hard lag in the 5–6-Hz QPO of Cyg X-2, but this phenomenon may not be due to Compton scattering, because of the large lag time.

Thus some of these phenomena suggest the existence of low-energy-gain (or loss) Compton scattering in these X-ray sources. None of them, however, suggests the existence of the high-energy-gain Compton scattering predicted in low-mass X-ray binaries by the model of White *et al.*¹⁹, which is similar to the inverse Compton scattering model of Cyg X-1 (refs. 3, 4).

It is important to search for these phenomena in other X-ray sources, so as to elucidate the cause of the time-lag behaviour and to investigate whether or not the kind of time delay shown in Fig. 1 is related to the presence of a black hole. For the same reason it would be interesting to search for these phenomena in active galactic nuclei and to compare them with those of Cyg X-1.

We wish to thank the members of the Ginga team, and also Dr Toshikazu Ebisuzaki for valuable discussions. We are grateful to Professor Walter H. G. Lewin and Mr Wim Penninx for valuable suggestions in preparing the manuscript, and Miss Seiko Mizobuchi for her help in calculations. This work was partly supported by a Grant-in-Aid for Scientific Research.

Occultation evidence for an atmosphere on Pluto

W. B. Hubbard*, D. M. Hunten*, S. W. Dieters†, K. M. Hill† & R. D. Watson†

* Lunar and Planetary Laboratory, University of Arizona, Tucson, Arizona 85721, USA

† Department of Physics, University of Tasmania, Hobart, Tasmania 7001, Australia

On 9 June 1988, Pluto occulted a 12th-magnitude star¹. Several observations of the occultation were obtained from Australia, New Zealand, and the south Pacific² and indicated that the initial decline in stellar flux was gradual, as would be expected if the starlight was defocused by an extended atmosphere around the planet. Here we interpret data obtained from the 1-m telescope at the University of Tasmania, Hobart, in terms of a theory for occultation by an atmosphere whose thickness is comparable to the planetary radius. The data can be satisfactorily fitted with a methane atmosphere at plausible pressures and temperatures. The surface pressures inferred from this single chord are uncertain by an order of magnitude, but are consistent with spectroscopic constraints.

To maximize the signal, observations were made in unfiltered light through a 30-arcsec aperture, including both Pluto and the occulted star. The detector was an EMI 9858A-S20 photomultiplier, and data were recorded at 1 kHz. At a later time, on 3 July 1988, it was possible to observe Pluto and the star separately with the same system at the same airmass, and the 9 June data were then corrected for the background from Pluto and sky, assuming that the regular variation of Pluto in the standard visual (V) filter (R. Marcialis, private communication) had the same relative amplitude in the unfiltered light. The data are shown in Fig. 1, in which ϕ represents the resulting normalized stellar signal, in units of the unocculted value, averaged to a time resolution of 1 s. The stellar signal displays a gradual variation characteristic of the presence of an atmosphere, rather than the abrupt change that would be expected from occultation of the star by an airless body.

At the time of the occultation at Hobart, one second of time corresponded to a projected sky-plane displacement of the star equal to 18.26 km with respect to Pluto's disk. Thus the total duration of the occultation, ~ 115 s, corresponds to a chord length at Hobart of $\sim 2,100$ km. This is comparable to the radius

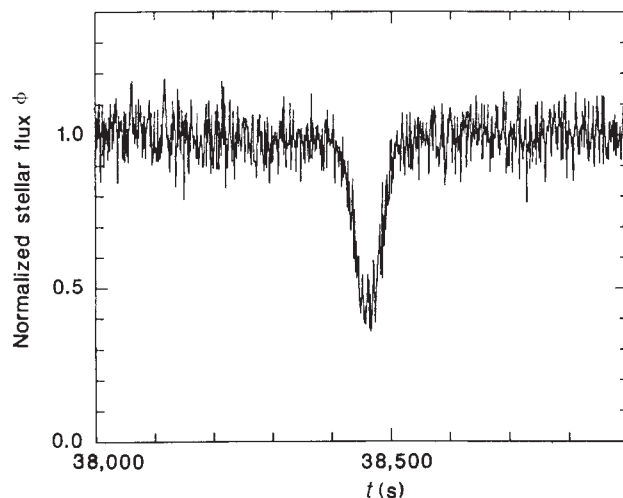


Fig. 1 Light curve observed at Hobart (t is seconds after 0:00:00 UT on 9 June 1988).

Received 28 March; accepted 6 October 1988.

- Oda, M. *Sp. Sci. Rev.* **20**, 757–813 (1977).
- Liang, E. P. & Nolan, P. L. *Sp. Sci. Rev.* **38**, 353–384 (1984).
- Shapiro, S. L., Lightman, A. P. & Eardley, D. M. *Astrophys. J.* **204**, 187–199 (1976).
- Sunyaev, R. A. & Trumper, J. *Nature* **279**, 506–508 (1979).
- Brinkman, A. C. *et al. Astrophys. J.* **188**, 603–608 (1974).
- Weiskopf, M. C., Kahn, S. M. & Sutherland, P. G. *Astrophys. J.* **199**, L147–L151 (1975).
- Priedhorsky, W. *et al. Astrophys. J.* **233**, 350–363 (1979).
- Nolan, P. L. *Astrophys. J.* **246**, 494–501 (1981).
- Page, C. G., Bennetts, A. J. & Ricketts, M. J. *Sp. Sci. Rev.* **30**, 369–371 (1981).
- Page, C. G. *Sp. Sci. Rev.* **40**, 387–390 (1985).
- Van der Klis, *et al. Astrophys. J.* **319**, L13–L18 (1987).
- Stella, L. *et al. Astrophys. J.* **327**, L13–L16 (1988).
- Makino, M. & the Astro-C team. *Astrophys. Lett. Commun.* **25**, 223–233 (1987).
- Bendat, J. S. & Piersol, A. G. *Random Data: Analysis and Measurement Procedures* 193 (Wiley, New York, 1971).
- Sunyaev, R. A. & Titarchuk, L. G. *Astr. Astrophys.* **86**, 121–138 (1980).
- Payne, D. G. *Astrophys. J.* **237**, 951–963 (1980).
- Stollman, G. M., Hasinger, G., Lewin, W. H. G., van der Klis, M. & van Paradijs, J. *Mon. Not. R. astr. Soc.* **227**, 7–12 (1987).
- Mitsuda, K. & the Ginga team in *The Physics of Compact Objects: Theory Versus Observations* (eds Filipov, L. & White, N. E.) (Pergamon, Oxford, in the press).
- White, N. E., Stella, L. & Parmar, A. N. *Astrophys. J.* **324**, 363–378 (1988).

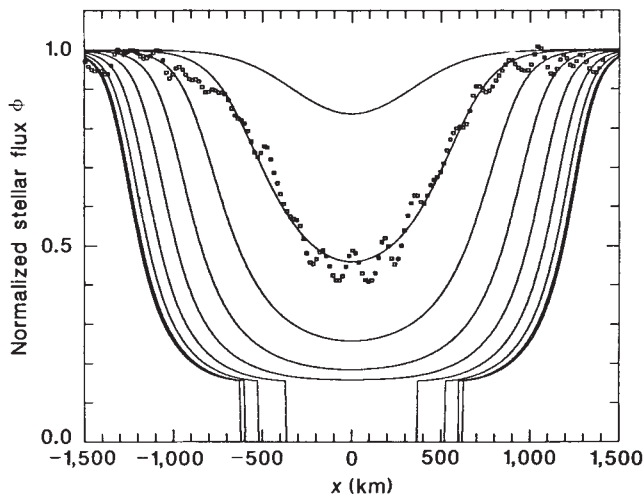


Fig. 2 Stellar flux plotted against position along chord, for nine chords in the warm atmosphere model. The chords shown range from $y_0 = 0$ –1,335 km, in steps of 167 km, with the Hobart chord at $y_0 = 1,168$ km (26 km outside the limb) being fitted by least squares. Squares show Hobart data, smoothed by a gaussian window with a 2-s characteristic width.

of the solid planet (D. J. Tholen, private communication), $a_0 = 1,142 \pm 21$ km. These numbers are more remarkable than they appear, because we find that Hobart was actually slightly outside the geometrical shadow of the solid planet. Evidently the Pluto atmosphere revealed by the occultation at Hobart is comparable in thickness to the radius of the planet itself, as has been suggested for some time^{3,4}. For this reason, the standard stellar occultation theory used for thin atmospheres^{5,6} requires modification.

For the initial analysis, we derive a theory for an isothermal atmosphere at temperature T , which is taken to be the skin temperature of Pluto's atmosphere, or $\sim 84\%$ of the surface temperature. Temperatures that fit the data are, as discussed below, comparable to the surface temperatures expected for Pluto. With a surface pressure of only a few microbars, comparable to that of the Earth's atmosphere at 80 km, standard ideas on thermal structure may not apply; the isothermal assumption is the simplest and should be adequate over the small height range that controls the light curve.

We assume that the Pluto atmosphere is composed of a single molecular species with molecular mass m , and define the parameter

$$\lambda = GMm/rk_B T \quad (1)$$

where G is the gravitational constant, M is Pluto's mass (which we take to be 1.29×10^{25} g), r is the distance from the centre of Pluto and k_B is Boltzmann's constant. Then the number density of molecules in Pluto's atmosphere in the region probed by the occultation is given by

$$n(r) = n_0 \exp(\lambda - \lambda_0) \quad (2)$$

where n_0 is the surface density and $\lambda_0 = \lambda(r = a_0)$. Equation (2) is valid in hydrostatic equilibrium, which is well satisfied in this problem, and $\lambda \gg 1$, which is also well satisfied for plausible molecules and values of T in the region probed by the occultation. Because both the star and observer are at large distances from Pluto compared with r , one may replace Pluto's spherical atmosphere with an equivalent two-dimensional phase-changing screen⁷ which changes the phase of an incoming plane wave from the star by an amount $\Phi(r)$, where r is now the distance from the centre of mass projected in the plane of the sky. Φ is calculated by integrating the refractivity of the gas, N , along a

straight path through the atmosphere with impact parameter equal to the distance r , and multiplying by the wavenumber of the starlight, k . Because N is proportional to n , the integrated phase shift produced by the atmosphere on a wavefront of starlight is⁸

$$\Phi = kN \left(1 + \frac{9}{8\lambda} \right) \sqrt{2\pi r^2/\lambda} \quad (3)$$

Equation (3) includes a first-order correction term in $1/\lambda$, which ultimately contributes a correction of about 1% to the occultation light-curve. All quantities are calculated at the radius of closest approach, that is, the projected distance from the centre of mass in the plane of the sky.

The total bending angle, α , of a ray after refraction by Pluto's atmosphere, with respect to its initial propagation direction, is

$$\alpha = -k^{-1} \frac{\partial \Phi}{\partial r}$$

We define r' as the observer's distance from the optical axis (defined by the line connecting the star and Pluto's centre), and D as the observer's distance from Pluto's centre (4.32×10^{14} cm). In general, for a spherical planet there will be two points where the total phase for a wavefront received by the observer is stationary: the near-limb point ($-$) and the far-limb point ($+$). We denote the values of r at these points as r_- and r_+ respectively. We have

$$D\alpha_- = r_- - r'$$

$$D\alpha_+ = r_+ + r'$$

respectively. For a given value of r' , these equations can be solved numerically for r_- and r_+ . The stellar flux is then calculated by obtaining the factors for one-dimensional defocusing in the radial direction:

$$\phi_{\pm} = \frac{1}{1 - D \frac{\partial \alpha_{\pm}}{\partial r_{\pm}}}$$

and then multiplying each by a factor that accounts for focusing in the orthogonal coordinate (which results from limb curvature). The total stellar flux, in units of the unocculted flux, is then given by⁹⁻¹¹

$$\phi = \frac{r_-}{r'} \phi_- + \frac{r_+}{r'} \phi_+$$

We neglect Fresnel diffraction by the planet's surface, and instead set $\phi_{\pm} = 0$ whenever $r_{\pm} < a_0$.

A likely candidate for the major atmospheric molecule of Pluto is methane, whose ice absorption features appear in the planet's spectrum^{12,13}. The maximum surface pressure of a methane atmosphere consistent with the spectroscopic evidence is 25 μ bar (refs 13, 14). The surface temperature of Pluto could plausibly vary between 46–64 K, depending on the rate of heating by sunlight and on thermal properties³. The corresponding vapour pressure of methane varies between much less than a microbar to ~ 1 millibar, which suggests that a tenuous methane atmosphere could be produced by sublimation of surface ice. No other plausible molecule has an appropriate vapour pressure in the expected temperature range. Sykes *et al.*¹⁴ find a sub-solar (presently equatorial) surface temperature of 58 ± 1 K, which corresponds to a methane vapour pressure of $\sim 80 \mu$ bar. The geometry of the occultation was such that the Hobart data pertain to an atmospheric region nearly above Pluto's equator.

We consider two models (cold and warm) which roughly bracket the expected temperature range for a Pluto methane atmosphere in the occultation region. The cold model assumes an atmosphere of pure methane at $T = 50$ K ($\lambda_0 = 28$), and the warm model assumes pure methane at $T = 61$ K ($\lambda_0 = 23$), close to the maximum surface temperature. The other parameter that

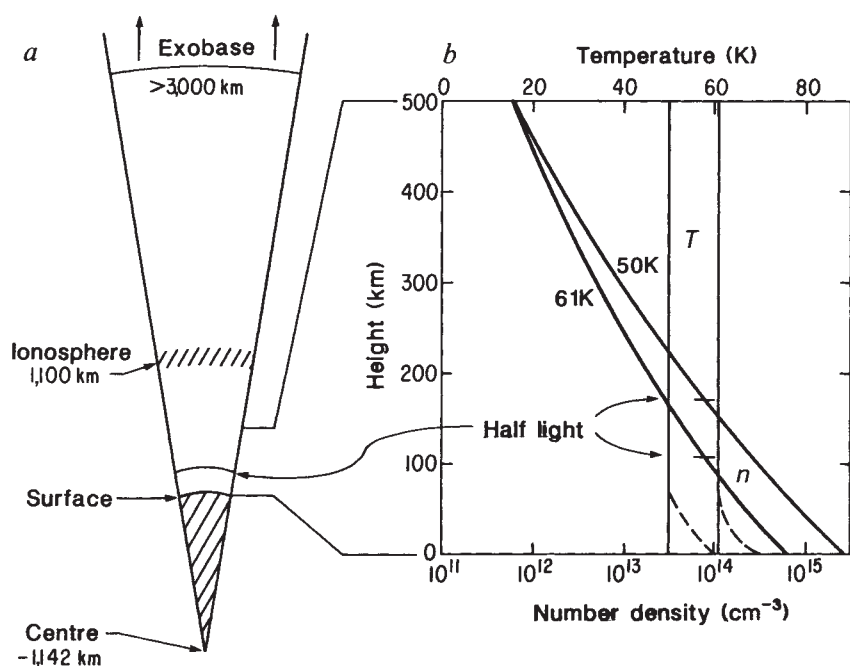


Fig. 3 A diagram of Pluto's atmosphere (a) and the warm (61 K) and cold (50 K) models (b). a Shows the main ray impact parameter at the level of half-light ($\phi = \frac{1}{2}$), the ionosphere and the exobase, above which molecules are essentially free of collisions and move on ballistic orbits. b Shows the temperature (T) and density (n) assumed for the bottom 500 km; the curvature of the density profile is due to variation of gravity with height. Dashed lines represent a possible modification of the isothermal profiles.

enters into the theory is the product $DN_0 = DN(r = a_0)$. Taking the refractivity of methane to be $N = N_{\text{STP}} = 4.41 \times 10^{-4}$ at STP¹⁵, we then have

$$DN_0 = 5.1 \times 10^7 \frac{P_0}{T}$$

where P_0 is the surface pressure in μbar , T is in K, and D is in cm. For convenience we write $DN = K_0 \exp(\lambda)$, where $K_0 \equiv DN_0$. Thus the parameters that are adjusted to fit the Hobart data are λ_0 , K_0 , y_0 (the minimum value of r') and t_0 (the time of closest approach at Hobart).

Figure 2 shows results of a fit of the warm model to the Hobart data. The parameter λ_0 was held fixed and K_0 , y_0 and t_0 were adjusted for best fit. In Fig. 2 we plot normalized flux ϕ as a function of position x along the chord (with $x = 0$ corresponding to closest approach). This model has $P_0 = 5 \mu\text{bar}$. As our data do not penetrate below the level $\phi \approx 0.46$, we do not observe the predicted rapid drop to $\phi = 0$ (total occultation) shown for more central chords (and which was observed from a station closer to the shadow's centre¹⁶), and we thus have no direct information about the location of a solid or opaque surface.

The cold model was fitted in a similar manner, yielding $y_0 = 1,234 \text{ km}$ and $P_0 = 18 \mu\text{bar}$. Although the grazing Hobart chord is fitted about as well by the cold model as by the warmer model, the central chords in the cold model exhibit a very different behaviour, because P_0 is sufficiently high that the star is never fully occulted by the surface. In contrast to the full stellar occultation for central chords predicted by the warm model, the cold model predicts that the stellar intensity should be enhanced by focusing near the centre of the shadow. The transition from a dark spot to a bright spot at the centre of Pluto's shadow occurs near $P_0 = 7 \mu\text{bar}$. The value of y_0 for the warm model is more consistent with a preliminary astrometric solution for the location of the Hobart chord with respect to the shadow's centre¹⁶, indicating that this is the more satisfactory of the two models.

Figure 3a shows a sketch of the very extended atmosphere for heights up to 3,000 km, and Fig. 3b shows the density models for $T = 50$ and 61 K. For $T = 50 \text{ K}$, the exobase is at 3,200 km, and the thermal escape flux, referred to the surface, is $4 \times$

$10^9 \text{ cm}^{-2} \text{ s}^{-1}$. Trafton³ has suggested that the escape flux might be much greater, enough to exhaust any reasonable supply of methane. But Hunten and Watson⁴ pointed out that such rapid escape would cool the upper atmosphere, so that the flux would in fact be limited to $\sim 5 \times 10^{10} \text{ cm}^{-2} \text{ s}^{-1}$. The limiting factor is the available solar energy at high altitudes, deposited in the region of the ionosphere (shown hatched in Fig. 3a). At $T = 61 \text{ K}$, the exobase is calculated to be well above 5,000 km, a sign that the atmosphere is on the verge of entering the state of energy-limited escape. Appreciable differences, as well as distortions from spherical symmetry, occur only at altitudes much higher than those probed by stellar occultation. Similarly, if there is a warmer region very near the surface, as indicated by the dashed lines in Fig. 3b, the effect should be small.

A combined analysis of all available data, together with coordinated observations of future Pluto occultations, should give substantial information about the atmosphere.

We thank R. L. Millis for predictions of the circumstances of this event, J. L. Elliot for a preprint reporting his observations, D. J. Tholen for new data on Pluto's radius, and R. Marcialis for help with determining the Pluto background. L. Carrasco, A. Klemola, R. P. Kraft, B. Weaver and R. J. Weymann assisted in plans to observe the occultation from North America (which were unrealized after the path became better known). This data analysis was supported in part by a NASA grant.

Received 11 July; accepted 18 October 1988.

1. Mink, D. J. & Klemola, A. *Astr. J.* **90**, 1894-1899 (1985).
2. Marsden, B. G. *IAU Circ.* nos 4611, 4612 (1988).
3. Trafton, L. *Icarus* **44**, 53-61 (1980).
4. Hunten, D. M. & Watson, A. J. *Icarus* **51**, 665-667 (1982).
5. Baum, W. A. & Code, A. D. *Astr. J.* **58**, 108-112 (1953).
6. Bartholdi, P. & Owen, F. *Astr. J.* **77**, 60-65 (1972).
7. Hubbard, W. B., Jokipii, J. R. & Wilking, B. A. *Icarus* **34**, 374-395 (1978).
8. Chamberlain, J. W. & Hunten, D. M. *Theory of Planetary Atmospheres: An Introduction to their Physics and Chemistry* 2nd edn (Academic, New York, 1987).
9. Young, P. J. *Astr. J.* **82**, 75-78 (1977).
10. Elliot, J. L. *et al. Astrophys. J.* **217**, 661-679 (1977).
11. Lellouch, E. *et al. Nature* **324**, 227-231 (1986).
12. Cruikshank, D. P., Pilcher, C. B. & Morrison, D. *Science* **194**, 835-837 (1976).
13. Buie, M. W. & Fink, U. *Icarus* **70**, 483-498 (1987).
14. Sykes, M. V., Cutri, R. M., Lebofsky, L. A. & Binzel, R. P. *Science* **237**, 1336-1340 (1987).
15. Allen, C. W. *Astrophysical Quantities* 3rd edn (Oxford University Press, 1973).
16. Elliot, J. L. *et al. Icarus* (submitted).

Formation of the Antarctic ozone hole by the ClO dimer mechanism

J. W. Barrett, P. M. Solomon, R. L. de Zafra,
M. Jaramillo, L. Emmons & A. Parrish*

Department of Earth and Space Sciences, State University of New York, Stony Brook, New York 11794-2100, USA

The discovery in 1986 of an extremely large concentration (~ 1 part per 10^9) of chlorine monoxide (ClO) at low altitudes in the spring stratosphere over Antarctica^{1,2}, and measurements showing the formation of OCIO at night³, provided strong evidence that the evolution of the Antarctic ozone 'hole' is chemically driven by chlorine. Here we use new measurements⁴ of the low-altitude ClO profile, made during September 1987, along with detailed observations of ozone depletion⁵ over McMurdo Station during the same period, to show that both the rate and altitude range of ozone depletion can be quantitatively accounted for by a mechanism in which the ClO dimer⁶ is the important intermediary in the catalytic destruction of ozone. An alternative bromine mechanism appears capable of contributing only 5–15% to the ozone loss rate.

All chlorine-driven mechanisms for ozone depletion have in common the reaction:



In the middle stratosphere this is turned into a catalytic cycle through reaction with atomic oxygen⁷:



This mechanism fails in the lower stratosphere, however, because of the extremely low O atom concentration, and a different explanation must be provided for the low-altitude ozone loss characteristic of Antarctica. The most viable mechanisms suggested as alternatives to reaction (2) in the Antarctic springtime lower stratosphere are the formation of ClO dimers⁶ and the combination of ClO with bromine oxide⁸. Both of these require prior heterogeneous chemistry on polar stratospheric clouds or other processes to remove NO_2 and alter the partition of chlorine, putting more of it into the active form. Evidence for denitrification is now quite strong^{9–12}. Moreover, both the ClO dimer and $\text{BrO} + \text{ClO}$ reactions, in addition to returning active chlorine to the destruction cycle, have rates that depend on the square of the ClO density or on the product of BrO and ClO densities, respectively, and can thus help to explain the rapid onset of the ozone hole over the past 10 years, during which Cl and Br have increased relatively slowly.

Our ground-based observations of ClO use a cooled heterodyne receiver operating at a frequency of 278,632 MHz to measure thermal emission from the $J = 15/2 \rightarrow 13/2$ rotational transition. Our measurements during 1986 clearly showed the presence of a layer of extremely high ClO concentration in the lower stratosphere at ~ 20 km (mixing ratio ~ 1.5 parts per 10^9 by volume (p.p.b.v.)). Because of the instrument's limited spectral bandpass of 256 MHz, we could not unambiguously obtain the shape of the altitude profile below 20 km. Our new measurements employed twice the bandwidth used in 1986, enabling us to measure the ClO concentration between 17 and 45 km. Observations were carried out during September and early October, and a full description including the diurnal cycle is presented elsewhere⁴. Here we concentrate on the average daytime ClO profile during the period with highest-quality data (20–24 September 1987). 'Daytime' is defined as beginning two hours after stratospheric dawn and ending two hours before stratospheric sunset.

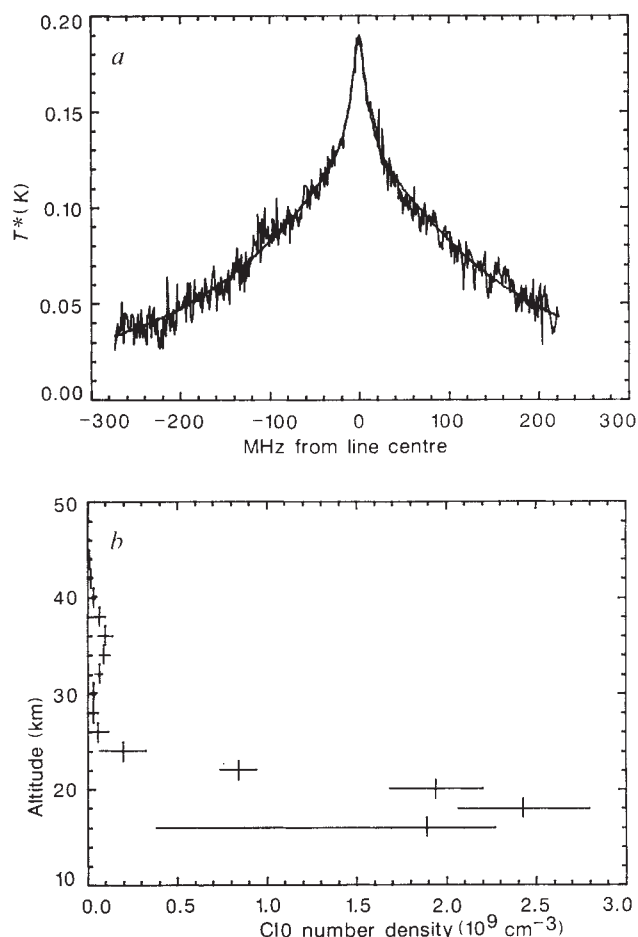


Fig. 1 *a*, Average daytime spectral lineshape of the 278,632-MHz rotational line of stratospheric ClO over McMurdo Station, Antarctica, 20–24 September 1987. The smooth solid line is the fit to the data obtained with the ClO concentration profile given in *b*. T^* is the Rayleigh-Jeans radiation temperature towards the zenith. *b*, Average daytime concentration profile of ClO over McMurdo Station, Antarctica, 20–24 September 1987, obtained by deconvolving the spectrum shown in *a*. The initial guess for this retrieval was a constant mixing ratio of 0.5 p.p.b.v. between 14 and 50 km. To estimate recovery errors, we have deconvolved a large number of synthetic spectra consisting of the noise-free lineshape calculated from a similar profile with randomly generated gaussian noise added at the same level as the noise in the data. The ensemble of recovered profiles was analysed statistically to obtain the 2σ (95% confidence level) error bars shown. Further error estimates were obtained, independently of the retrieval process, by varying the ClO concentration at a particular altitude and performing χ^2 tests on the difference between the forward calculated spectrum and the data. These tests showed that below 17 km the retrieval accuracy was much less than indicated by the Monte Carlo tests and limited by the bandwidth of the observations. The errors in the 15–17 km range reflect the results of the χ^2 tests. An additional instrumental calibration uncertainty of $\pm 12\%$, independent of altitude, has not been included in these error bars.

Figure 1a shows the spectral line obtained during this period; the extremely large quantity of low-altitude ClO present is immediately noticeable from the pressure-broadened line wings and from the total integrated intensity under the spectrum, which is twenty times that of a typical mid-latitude observation. The altitude profile is obtained from the shape and intensity of the pressure-broadened spectral line, using a modified Chahine-Twomey deconvolution algorithm¹³. For the initial guess of the altitude profile (required by the algorithm) we used a constant 0.5 p.p.b.v. above 15 km. The retrieved altitude profile between

* Also at Millitech Corp., S. Deerfield, Massachusetts, USA.

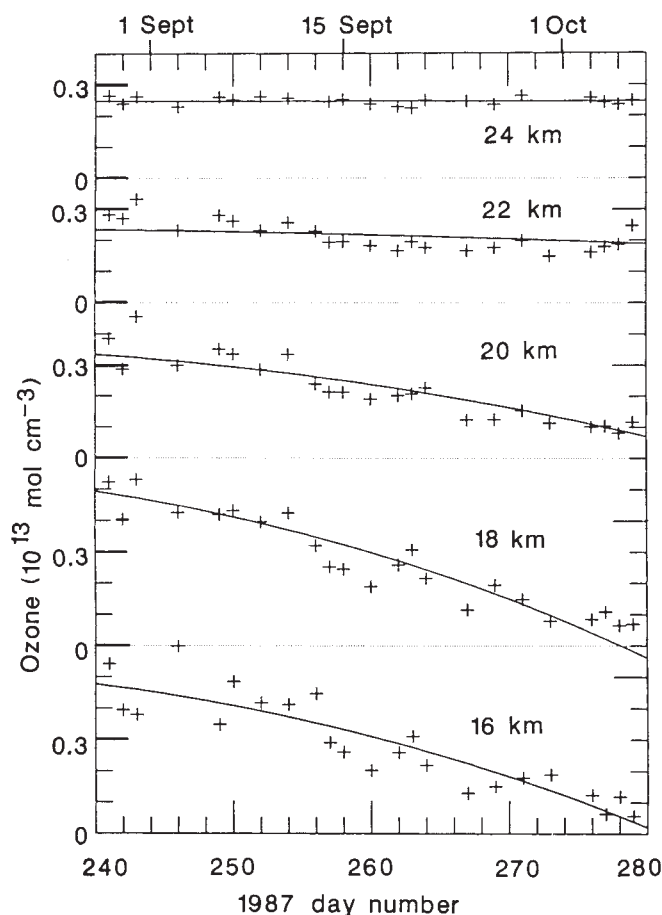
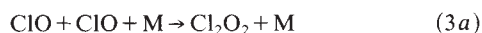


Fig. 2 Decline of ozone concentration over McMurdo Station, Antarctica, during the 1987 Austral spring as a function of altitude. The observed points (+), are from balloon measurements of the ozone profile by Hofmann *et al.*⁵. The solid line is the decline in ozone expected from the ClO dimer mechanism using our measurements of ClO (see equations (4) and (5)). Both the rate of decline and the altitude dependence, particularly the fall-off in depletion above 21 km, are well matched by the dimer depletion rates from the ClO measurements. We have included the lowest measured altitude bin at 15–17 km for completeness, although the errors in the ClO concentration are much larger than those above 17 km.

17 and 27 km changes by less than $\pm 10\%$ for initial guesses in the range 0.1–1.0 p.p.b.v. Figure 1b shows that the derived average daytime ClO number density profile is bimodal, with maxima at 18 km and at ~ 36 km. Here we shall be concerned only with the (dominant) lower component.

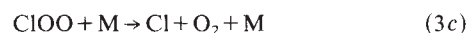
In the lower stratosphere, the average daytime ClO mixing ratio is 1.3 ± 0.2 p.p.b.v., centred at an altitude of 19.5 ± 1 km, falling to half peak at 16 and 23 km, and falling below 0.3 p.p.b.v. above 24 km. Thus, during the day, about half of the total chlorine is in the active form, ClO, based on an assumed total chlorine mixing ratio of 3 p.p.b.v. by 1987 (ref. 14). For determining depletion rates it is preferable to consider the ClO density profile in Fig. 1b. This shows a layer with a peak density of 2.4×10^9 cm⁻³ at 18 km, extending from 22 km down to ≤ 16 km; this layer corresponds very closely to the altitude range of ozone depletion found above McMurdo during the same time period⁵.

These very high ClO densities favour the dimer-driven depletion mechanism⁶, which depends on the square of the ClO density. This involves ozone destruction by chlorine (reaction (1)), followed by dimer formation:



and processes that return two free chlorine atoms from the dimer

to close the catalytic cycle:

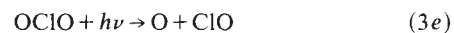


We compare the expected rate of ozone removal by the dimer mechanism (reactions (1) and (3a–c)), using observed ClO concentrations, with the actual depletion rate observed as a function of altitude during September 1987, at $\sim 78^\circ$ S.

A key question in this process is whether other dimer decay channels exist that can return ClO in any appreciable quantities from the dimer and reduce the net ozone depletion. These alternative channels, which result in no net ozone depletion, include the back reaction to (3a), as well as:



followed by



Cox and Hayman¹⁵ find that the reverse reaction to (3a) is very slow (~ 10 hours) at $T \leq 200$ K, whereas photolysis of the dimer requires < 1 hour in the daylight stratosphere. Our own observation⁴ that the low-altitude ClO concentration drops by more than a factor of ten by sunset and remains low throughout the night also implies a very slow back-reaction rate, and the rapid post-dawn rise in ClO implies a fast photolysis rate. Cox and Hayman also conclude that significant OCIO is not produced by photolysis (reactions (3d) and (3e)). Field observations, showing that OCIO is less concentrated during morning twilight than during evening twilight, reinforce this conclusion¹⁶.

In the following we assume minimal horizontal transport of ozone from lower latitudes during the highly stable 1987 vortex, and negligible vertical transport in view of the aerosol layer stability observed by Hofmann *et al.*¹⁷. There is no significant *in situ* production of ozone by solar ultraviolet radiation at this latitude, altitude and season. Thus, ozone change should be dominated by the depletion mechanism.

As long as $[\text{O}_3] \gg [\text{ClO}]$, the rate-limiting step is formation of the ClO dimer (reaction 3(a)). We assume that two free chlorine atoms, each of which will destroy one ozone molecule, are released for each Cl_2O_2 formed. Thus, the rate of ozone destruction will be twice the rate of formation of the ClO dimer:

$$\frac{d}{dt}[\text{O}_3] = -2k_{3a}[\text{ClO}]^2[\text{M}] \quad (4)$$

Hayman *et al.*¹⁸ have determined k_{3a} to be $6.0 \pm 0.3 \times 10^{-32}$ ($T/300$)^{-2.1 $\pm$ 0.7} cm⁶ s⁻¹. In using equation (4) we will use only observed daytime ClO concentration and air density [M] as a function of altitude, so that no knowledge is required of the actual release rate for two chlorines from the dimer. Note that the dependence on the background gas concentration [M] provides a mechanism for sustaining substantial ozone depletion well below the peak of the ClO mixing-ratio profile, even though ClO concentration is then low; this is not true for the bromine mechanism.

For continuous daylight, the rate of ozone depletion predicted from equation (4) would be constant if [ClO] remained unchanged, limited by the available active chlorine and its partitioning. We have experimental evidence² that this is approximately the case throughout the period of greatest ozone depletion, from late August to late September. For simplicity we shall treat daytime [ClO] as constant, equal to the value during 20–24 September. Measurements of column NO₂ over McMurdo¹⁹ in 1987 show little or no increase in average column over days 247–275, followed by a sharp increase beginning at about day 280, which we take as indirect evidence supporting the assumption of constant [ClO]. The amount of sunlight per calendar day at an altitude of around 20 km varies greatly, however, from ~ 9 hours on 1 September to 18 hours on 30

Table 1 Daytime ozone depletion rates over McMurdo Station, Antarctica, September 1987

Altitude (km)	ClO mixing ratio* (p.p.b.v.)	[ClO] (10 ⁹ cm ⁻³)	$\frac{d}{dt}[\text{O}_3](10^6 \text{ cm}^{-3} \text{ s}^{-1})$	
			Observed†	Calculated (dimer mechanism with observed ClO)*
15–17	0.6 ^{+0.09} _{-0.25}	1.8	-3.4 ± 0.4	-3.3 ^{+2.1} _{-1.0}
17–19	1.1 ± 0.15	2.3	-3.4 ± 0.3	-3.7 ± 1.0
19–21	1.3 ± 0.18	1.9	-2.2 ± 0.2	-1.8 ± 0.5
21–23	0.8 ± 0.11	0.9	-0.74 ± 0.18	-0.3 ± 0.1
23–25	0.3 ± 0.09	0.2	-0.05 ± 0.06	-0.01 ± 0.01

* One standard deviation errors from the calibration of the millimetre-wave ClO measurements and profile retrieval. Percentage errors in [ClO] are the same as in the mixing ratio. Errors from the uncertainty in k_{3a} (which affect the calculated ozone depletion rate from the ClO dimer mechanism) are discussed in the text.

† The observed ozone depletion rate and associated 1 σ errors are from the least-squares fit of equation (5) to the ozone data of Hofmann *et al.*⁵.

September. The observed diurnal cycle for ClO (refs 1, 4) indicates that its concentration remains relatively constant during the 'daytime', and we find that the relevant average for use in equation (4), $\langle[\text{ClO}]^2\rangle$, is close to $\langle[\text{ClO}]\rangle^2$. ClO decreases at least tenfold by sunset, signifying the end of active ozone destruction until the following dawn. If t_A represents the cumulative hours of daylight during which the dimer cycle has been active after day 240 (28 August) then the ozone concentration remaining by day D will be

$$\text{O}_3(D) = \text{O}_3(D=240) - \frac{d}{dt}[\text{O}_3] \times t_A(D) \quad (5)$$

Equations (4) and (5), along with measured values of $[\text{O}_3]$ on day 240, and knowledge of $[M]$, T , k_{3a} and $[\text{ClO}]$ as a function of altitude, allow depletion of ozone to be predicted as a function of altitude. These predictions can be compared with the ozone measurements of Hofmann *et al.*⁵ over McMurdo Station in 1987. They found a rapid and continuous depletion throughout the period covering days 240–280 (28 August to 7 October) in the altitude range from 15 to 21 km (see Fig. 2). Above 21 km there was little depletion after day 258, and above 23 km, no measurable depletion during the entire period. The observed ozone and ClO measurements represent the average conditions in the vortex at 78°S, because McMurdo was well within the region of major ozone depletion during the entire period, and the circumpolar transit time is only ~6 days.

In Table 1 we calculate $d[\text{O}_3]/dt$ from equation (4) as a function of altitude, using our measured [ClO]. To compare these predicted rates with observations, we have fitted equation (5) to the ozone data⁵ to obtain an 'observed' $d[\text{O}_3]/dt$ at each altitude range. These observed rates are also given in Table 1. Within experimental uncertainties, the observed and predicted depletion rates agree at all altitudes. Both observed and predicted rates decrease steeply at 22 km and are essentially zero at 24 km. Between 17 and 21 km, corresponding to the region of highest depletion rates, and where the measured ClO densities are most accurate, we find an average predicted depletion of $-2.7 \pm 0.5 \times 10^6 \text{ cm}^{-3} \text{ s}^{-1}$ and an essentially equal average observed depletion of $-2.8 \pm 0.2 \times 10^6 \text{ cm}^{-3} \text{ s}^{-1}$ (although there is significant uncertainty (~37%) in the laboratory measurement of k_{3a}).

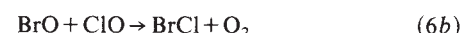
Figure 2 compares the decrease in ozone predicted by the dimer mechanism between days 240 and 280 with the data of Hofmann *et al.*⁵. Agreement is excellent at 18, 20 and 24 km; at 22 km the break at day 258 is not reproduced but the average observed and predicted depletion are both small. The ClO dimer mechanism also predicts no depletion where none is observed. Similar conclusions on the efficacy of the dimer mechanism below 18.5 km have been derived independently using aircraft measurements¹⁹. Note that the ozone data is consistent with a flattening of depletion rates at all altitudes after ~days 275–280,

resulting from the return of NO_2 (ref. 20) and a decline in ClO through recombination with NO_2 .

We now consider the possible role of the alternative ($\text{BrO} + \text{ClO}$) mechanism⁸. The rate-limiting step here is either



or



Measurements of stratospheric BrO over McMurdo²¹, made at the same time as our ClO measurements, indicated 5 to 10 parts per 10¹² by volume (p.p.t.v.) of BrO during days 262–268 (although the higher optical absorption cross-section reported by Sander and Watson²² would reduce these values by 30%). This is consistent with aircraft measurements²³ of BrO, which gave ~6 p.p.t.v. at ~18 km. Adopting a mean of 7 p.p.t.v. for the mixing ratio of BrO, and using our ClO measurements, we can calculate the depletion rates expected from the bromine cycle. There are two reported measurements of the combined rate constant for reactions (6a) and (6b): $k_6 = 3.7 \times 10^{-12} \text{ cm}^3 \text{ s}^{-1}$ (with no T dependence; ref. 24) and $k_6 = 10.5 \times 10^{-12} \text{ cm}^3 \text{ s}^{-1}$ (extrapolated to 195 K; ref. 25). Adopting the higher value, we calculate depletion rates $d[\text{O}_3]/dt$ of 0.5×10^6 , 0.4×10^6 and $0.25 \times 10^6 \text{ cm}^{-3} \text{ s}^{-1}$ at 16, 18 and 20 km respectively. Using the smaller rate constant would reduce these values by an additional factor of three, yet they are already a factor of seven or more below observed rates of depletion in Table 1.

We have demonstrated that the dimer mechanism is able to account for the observed rates and altitude range of ozone depletion using our observed concentration of ClO in the Antarctic spring stratosphere and the current measured value for the dimer formation rate constant. We conclude that bromine plays a relatively minor role in Antarctic ozone depletion.

This work was supported by NASA, the Chemical Manufacturers Association, and received generous support from the NSF's Division of Polar Programs. We especially thank D. J. Hofmann for making available his group's ozone measurements before publication.

Received 14 July; accepted 11 October 1988.

- De Zafra, R. L. *et al.* *Nature* **328**, 408–411 (1987).
- Solomon, P. M. *et al.* *Nature* **328**, 411–413 (1987).
- Solomon, S., Mount, G. H., Sanders, R. W. & Schmeltekopf, A. L. *J. geophys. Res.* **92**, 8329–8338 (1987).
- De Zafra, R. L. *et al.* *J. geophys. Res.* (in the press).
- Hofmann, D. J., Harder, J. W., Rosen, J. M., Hereford, J. V. & Carpenter, J. R. *J. geophys. Res.* (in the press).
- Molina, L. T. & Molina, M. J. *J. phys. Chem.* **91**, 433–436 (1987).
- Molina, M. & Rowland, F. S. *Nature* **249**, 810–814 (1974).
- McElroy, M. B., Salawitch, R. J., Wofsy, S. F. & Logan, J. A. *Nature* **321**, 759–762 (1986).
- McKenzie, R. L. & Johnson, P. V. *Geophys. Res. Lett.* **11**, 73–75 (1984).
- McCormick, M. P. & Larsen, J. C. *Geophys. Res. Lett.* **13**, 1280–1283 (1986).
- Mount, G. H., Sanders, R. W., Schmeltekopf, A. L. & Solomon, S. *J. geophys. Res.* **92**, 8320–8328 (1987).

12. Farmer, C. B., Toon, G. C., Schaper, P. W., Blair, J.-F. & Lowes, L. L. *Nature* **329**, 126–130 (1987).
13. Twomey, S., Herman, B. & Rabinoff, R. J. *atmos. Sci.* **34**, 1085–1090 (1977).
14. *Atmospheric Ozone 1985* (Rep. No. 16 of the WMO Global Ozone Research and Monitoring Project) (World Meteorological Organization, 1986).
15. Cox, R. A. & Hayman, G. D. *Nature* **332**, 796–800 (1988).
16. Solomon, S., Sanders, R. W., Carroll, M. A. & Schmeltekopf, A. I. *J. geophys. Res.* (in the press).
17. Hofmann, D. J., Rosen, J. M., Harder, J. W. & Hereford, J. V. *J. geophys. Res.* (in the press).
18. Hayman, G. D., Davies, J. M. & Cox, R. A. *Geophys. Res. Lett.* **13**, 1347–1350 (1986).
19. Anderson, J. G. *et al. J. geophys. Res.* (in the press).
20. Sanders, R. W., Solomon, S., Carroll, M. A. & Schmeltekopf, A. I. *J. geophys. Res.* (in the press).
21. Carroll, M. A., Sanders, R. W., Solomon, S. & Schmeltekopf, A. I. *J. geophys. Res.* (in the press).
22. Sander, S. P. & Watson, R. T. *J. phys. Chem.* **85**, 4000–4007 (1981).
23. Brune, W. H. & Anderson, J. G. *J. geophys. Res.* (in the press).
24. Hills, A. J., Cicerone, R. J., Calvert, J. G. & Birks, J. W. *Nature* **328**, 405–408 (1987).
25. Sander, S. P. & Friedl, R. R. *Geophys. Res. Lett.* **15**, 887–890 (1988).

Standing and propagating wave oscillations in the anodic dissolution of nickel

O. Lev*, M. Sheintuch, L. M. Pisemen & Ch. Yarnitzky†

Department of Chemical Engineering, and

† Department of Chemistry,
Technion-Israel Institute of Technology,
Haifa 32000, Israel

Spontaneous pattern formation in reactive media has been thoroughly investigated in homogeneous oscillatory systems, mainly in the Belousov–Zhabotinsky reaction¹. Spatiotemporal patterns have been observed in some studies of catalytic surfaces^{2,3} and of travelling electrochemical waves⁴. Here we report on oscillatory wave patterns observed during electrochemical dissolution of a nickel wire in acidic media. We show that space-averaged potential or current oscillations are associated with the creation of an inhomogeneous current distribution, and that the selection of a specific spatial current pattern depends on the current control mode of the electrochemical cell. In the almost potentiostatic (fixed potential) mode of operation, a train of travelling pulses prevails, whereas antiphase oscillations occur in the galvanostatic (constant average current) mode. The latter, as far as we know, have never before been reported in any reactive system.

Travelling electrochemical waves were reported at the turn of the century by Heathcote⁴, who observed activity and passivity fronts propagating in the dissolution of iron in nitric acid solution. This phenomenon served for years as a model for axon excitation and information transfer in biological systems. Similar effects have been observed during other electrochemical dissolution reactions. Lumped (space-averaged) galvanostatic oscillations in the anodic dissolution of nickel in sulphuric acid solution were observed in the 'transpassive' region^{5,6}. This is a region in the current–potential diagram that is characterized by dissolution of the nickel either by diffusion through the oxide layer or by localized oxidation through patches in the oxide film. In the potentiostatic mode of operation (that is, at a fixed potential difference between the active electrode and a reference electrode), a unique stable solution was found to prevail throughout the transpassive branch. Osterwald⁵ noted that inserting an external electric resistance between the two electrodes, at a fixed potential difference, creates an intermediate mode of operation between the potentiostatic and galvanostatic ones and may yield oscillatory behaviour. We have studied the spatial inhomogeneities during galvanostatic and almost potentiostatic oscillations.

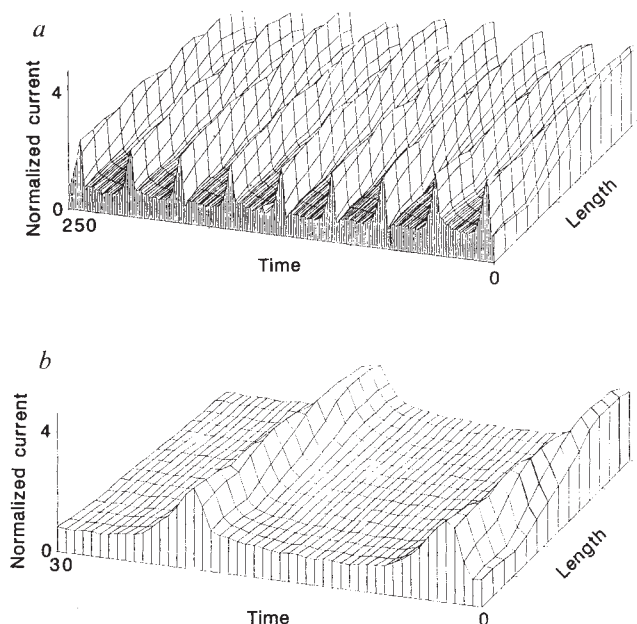


Fig. 1 Travelling waves during almost potentiostatic oscillations: *a*, Three-dimensional representation of local current distribution. *b*, Enlarged section. Operating conditions: 20-cm nickel wire in 6N sulfuric acid, 1.67 V relative to saturated Calomel electrode, current normalized relative to 65 mA cm^{-2} , distributed homogeneously. Timescale in scans (50 scans s^{-1}); length scale covers the whole wire.

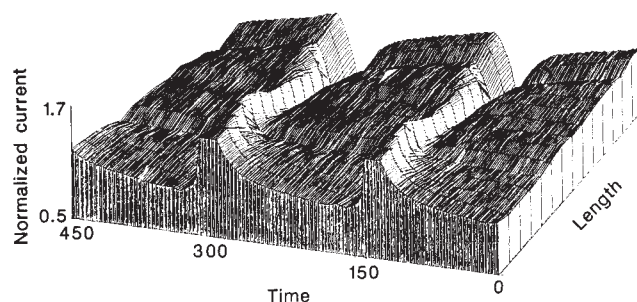


Fig. 2 Local current distribution of antiphase oscillations during galvanostatic operation. Operating conditions: 20-cm nickel wire in 2N sulfuric acid, current is 16.25 mA cm^{-2} ; local current normalized relative to the current density at 15.5 mA cm^{-2} as measured under potentiostatic conditions ($I = (v_g - v_0)/(v_p - v_0)$, where v_0 , v_g and v_p are the micro-electrode potential at rest, under galvanostatic and under potentiostatic conditions, respectively). Time and length scales are the same as in Fig. 1.

The principle of the experiments is as follows. The current flux between two parallel electrodes immersed in shallow electrolytic solution is essentially perpendicular to the plane of these electrodes. The potential of micro-reference electrodes, situated sufficiently close to the working electrode, provides a direct measure of the local current density. This holds provided that the potential of this electrode is recorded relative to a common electrode located in a position that is insensitive to the current between the main two electrodes.

The electrochemical cell ($200 \times 200 \times 40 \text{ mm}$) is constructed of plexiglass and holds two parallel nickel wires. The wires are 99.8% pure nickel (Alfa Products), 0.63 mm in diameter and 200 mm in length, and are situated 12 mm apart. A Calomel reference electrode is situated 40 mm behind the counter electrode at the centre of the cell. The cell is controlled by a galvanostat and a potentiostat in parallel, with a switch to enable

* Present address: Division of Human Environmental Sciences, Graduate School of Applied Science and Technology, The Hebrew University of Jerusalem.

rapid exchange of the operational modes. The current distribution is recorded by fifteen Ag/AgCl micro-reference electrodes equally spaced along the working electrode relative to a common Calomel electrode situated near the cell reference electrode.

A typical train of waves observed during almost potentiostatic conditions in the oscillatory domain is shown in Fig. 1. An external 1- Ω resistor was used to convert the stationary homogeneous current distribution into a sequence of pulses. The pulses emanated from one end of the working electrode (right-hand side of Fig. 1) and travelled at $\sim 4.5 \text{ ms}^{-1}$ to the other end of the electrode.

Typical antiphase oscillations observed during galvanostatic operation are shown in Fig. 2. There is a phase shift of π between the opposite oscillations at electrode edges. This phenomenon persists for nickel wires of various size (1–20 cm long) and over a wide range of acid concentrations (1–6 M). Coherent spatial organization occurs even when the temporal behaviour is aperiodic.

Chemical oscillations differ from hydrodynamic ones in that the former can be observed in a completely mixed (homogeneous) system, whereas the latter result from spatial inhomogeneity. Thus, the simplest time-dependent solution to be expected in chemical systems involves homogeneous oscillations. The problem of synchronization along chemical, and specifically heterogeneous, reacting systems, and the existence of inhomogeneous states, is strongly reminiscent of similar phenomena in tissues of cells, an analogy which dates back to Turing's pioneering work on morphogenesis⁷.

The pattern observed in the almost potentiostatic mode of operation (Fig. 1) shows good global synchronization among the local oscillators. Such a pattern is expected in diffusion-

reaction systems with oscillatory kinetics, and similar patterns were observed in the Belousov-Zhabotinsky reaction⁸. Pattern observations in the latter system, as well as in neural transmission, are also attributed to diffusion and reaction in excitable media⁸.

The condition of constant space-averaged current in the galvanostatic mode leads to antiphase oscillations. Because the local current oscillates, the wire edges must lie out of phase so that their sum will eliminate the average current amplitude. This type of operation has analogues in other chemically reacting systems subject to control, such as a catalytic wire operated in the constant resistance mode⁹. We believe that such symmetry-breaking causes many of the complex and chaotic motions observed in electrochemical and catalytic reactions. The possibility of such almost instantaneous synchronization modes occurring in nervous systems is intriguing and of great potential interest to the study of neurotransmission.

This work was supported by the US-Israel Binational Science Foundation.

Received 20 June; accepted 25 October, 1988.

1. Noszticzzi, Z., Horsthemke, W., McCormick, W. D., Swinney, H. L. & Tam, W. Y. *Nature* **329**, 619–620 (1987).
2. Brown, J. R., D'Netto, G. A. & Schmitz, R. A. in *Temporal Order*, Springer Series in Synergetics, Vol. 30 (eds Rensing, L. & Jaeger, N. I.) 86–95 (Springer-Verlag, Berlin, 1985).
3. Cox, M. P., Ertl, G. & Imbihi, R. *Phys. Rev. Lett.* **54**, 1725–1728 (1985).
4. Heathcote, H. L. *J. Soc. Chem. Ind.* **26**, 809–841 (1907).
5. Osterwald, J. & Feller, H. G. *J. Electrochem. Soc.* **105**, 473–XXX (1960).
6. Lev, O., Wolffberg, A., Sheintuch, M. & Pismen, L. M. *Chem. Engng Sci.* **43**, 1339–1353 (1988).
7. Turing, A. M. *Phil. Trans. R. Soc. B* **237**, 37–52 (1952).
8. *Oscillations and Travelling Waves in Chemical Systems* (eds Field, R. J. & Burger, M.) (Wiley, New York, 1985).
9. Sheintuch, M. & Schmidt, J. *Chem. Engng Commun.* **44**, 33–52 (1986).

Mantle metasomatism by ephemeral carbonatite melts

David H. Green & Margaret E. Wallace

Department of Geology, University of Tasmania, GPO Box 252C, Hobart 7001, Australia

Recent experiments¹ have shown that, at a range of upper-mantle temperatures and pressures, a carbonatite (carbonate-rich) melt occurs as a very small melt fraction in equilibrium with pargasite lherzolite. This melt has low water content and low TiO₂ content, and will be extremely effective in transporting large-ion lithophile elements. Here we show that such melts will alter the chemical composition of ('metasomatize') the mantle through which they pass, producing distinctive mineralogical and geochemical signatures. The predicted signatures can be recognized in peridotite xenolith suites from Australia, suggesting that metasomatism by carbonatite melts is an effective process for redistributing certain elements in the lithosphere, particularly in continental regions.

The chemistry of the Earth's upper mantle may be investigated by two methods: by studying high-pressure xenoliths in volcanic vents, lava flows or feeder dykes, or by studying mantle-derived magmas representing small melt fractions of the unknown mantle source. Both approaches have established that the Earth's mantle is mineralogically, chemically and isotopically heterogeneous on the hand-specimen scale, as illustrated by xenoliths^{2–5}, and on the large scale, as illustrated by isotopic and trace-element variations within and between volcanic provinces^{5–8}. In seeking to understand this heterogeneity, it is common to appeal to 'mantle metasomatism'⁹, a term used for all additions or removals of chemical components by means of melts or fluids, or by diffusive exchange.

Experimental studies of the peridotite-H₂O system^{10,11} constrained models of mantle metasomatism by hydrous silicate

melts which may be very enriched in incompatible elements (large-ion lithophile elements (LILE) and high-field-strength elements (HFSE)) in the mantle environment. The *P-T* stability relationships of pargasitic amphibole in 'fertile peridotite' are such that metasomatizing melts and/or aqueous fluids may be 'fixed' in the lithosphere at depths near the high-pressure limit of amphibole stability (25–30 kbar or 80–100 km)^{2,10–12}. Using as a basis an earlier experimental study of the peridotite-H₂O-CO₂ solidus¹³ in which the solidus of amphibole + carbonate lherzolite was interpreted as coincident with the disappearance of amphibole, recent experimental studies^{14,15} have explored in more detail the nature of metasomatism by carbonated silicate melts. Reaction of carbonated nephelinitic melts with residual mantle harzburgite or lherzolite may produce clinopyroxene-rich wall-rock assemblages of wehrlite or olivine clinopyroxenite at *P* < 17 kbar (ref. 14), leading to enrichment in alkali elements and CO₂ in the silicate magma, or to complete reaction to lherzolite with release of CO₂-rich fluid¹⁴. At higher pressures, similar reactions between carbonated silicate melt and harzburgite or lherzolite are inferred to produce amphibole + carbonate-bearing lherzolite¹⁵. Because of the interpretation that the solidus of carbonate + amphibole lherzolite coincides with amphibole disappearance and that initial melts of carbonate + amphibole lherzolite are carbonated silicate melts^{13–15}, with <10% CO₂ in solution, these models of metasomatism show metasomatic enrichment in the HFSE such as Ti and also the LILE such as Na and K.

Our current studies of the peridotite-CO₂-H₂O solidus¹ for small water and carbonate contents (Fig. 1) have revealed a *P-T* field in which a sodic, dolomitic carbonatite melt, with very low H₂O content, coexists with pargasite-bearing lherzolite. Our experimental data define a distinctive *P-T* window, which includes the conditions traversed by normal oceanic or young continental lithosphere geotherms (Fig. 1), in which a sodic, dolomitic carbonatite melt may occur and may cause mantle metasomatism (21–31 kbar, 930–1,080 °C for fertile lherzolite).

Other authors have argued a metasomatic role for ($\text{H}_2\text{O} + \text{CO}_2$) fluids¹⁶⁻²¹, estimating solute contents of <2 wt% for CO_2 -bearing fluids in peridotitic assemblages¹⁶. However, the carbonatite melt¹, with a CO_2 content of >40 wt% and an H_2O content estimated at 1-2%, is a different and far more effective agent for metasomatism. The low water content of the carbonatite melt is particularly significant and results from stable persistence of residual pargasitic amphibole, with water partitioning between the amphibole and the carbonatite melt.

The carbonatite melt is very effective in fractionating the LILE and phosphorus (partitioned into the carbonatite melt) from the HFSE (Ti has a low solubility in the carbonatite melt and is partitioned into pargasite, clinopyroxene and residual ilmenite). Our techniques do not permit trace-element analysis of the segregated carbonatite melt, but we may use trace-element contents of natural carbonatites to predict that the melt will be greatly enriched in Na, K, Sr, Rb and light rare-earth elements (LREE) and will be effective in quite specific trace-element fractionation, for example changes in Na/K, Rb/Sr, Sm/Nd, Nb/Ti and other element ratios among the LILE and HFSE groups. Metasomatism of lithospheric mantle by even very small amounts (<5%) of sodic, dolomitic carbonatite melts would dominate LILE and LREE abundances originally present in residual, refractory lherzolite or harzburgite in the lithosphere (compare components B and A respectively in ref. 2).

If melt segregation and aggregation occur within the carbonatite melt window, then this melt may move rapidly to the surface, by magma fracturing mechanisms, along narrow dykes or veins, yielding primary sodic, dolomitic carbonatite magmas or fractionated natro-carbonatite derivatives^{1,22-24}.

But if sodic, dolomitic carbonatite melts infiltrate lherzolite at pressures <21 kbar at 950-1,050 °C, then decarbonation reactions occur with release of CO_2 vapour. Simple endmember reactions may be written that collectively react alkali-element contents of carbonatite melt with alumina contents of orthopyroxene or spinel to form clinopyroxene and pargasite components, and react enstatite with dolomitic melt to yield diopside and olivine with CO_2 release. Collectively these reactions replace enstatite in lherzolite or harzburgite with clinopyroxene + olivine ± chrome spinel, that is, they modally alter lherzolite or harzburgite towards wehrlite compositions. The stability of amphibole and phlogopite will depend sensitively on both composition and CO_2 : H_2O ratios of the fluid released by decarbonation of the carbonatite melt.

The effect of the decarbonation reactions and release of CO_2 is to raise the solidus of the lherzolite either to the dehydration solidus for pargasite lherzolite at high f_{CO_2} (Fig. 1), or, if pargasite breaks down owing to high f_{CO_2} , then the solidus will be that for anhydrous peridotite or phlogopite peridotite at high f_{CO_2} , low $f_{\text{H}_2\text{O}}$. In addition to the experimental study of dolomite + pargasite peridotite¹, we have experimentally confirmed the decarbonation and solidus reactions of Fig. 1 by holding mixtures of peridotite + carbonatite melt (pargasite-bearing pyroxene + 2% sodic dolomitic carbonatite¹) at 18 kbar, 1,000 °C, for 4 h. The resulting charge consisted of olivine + enstatite + diopside_{ss} + spinel + pargasite + ilmenite (Table 1). There was no carbonate, carbonatite melt or silicate melt present and mineral compositions (Table 1) show the entry of calcium and alkali elements into pargasite and diopside and a decrease in aluminium content of orthopyroxene. Fluid inclusions were abundant, and particularly evident within olivine. The same mixture held at 22 kbar, 900 °C, for 26 h crystallized to dolomite + pargasite lherzolite (Table 1). The experiments at $P < 21$ kbar demonstrate the way that primary carbonatite melts may react to yield lherzolites or wehrlites in which the infiltration of carbonatite melt is revealed only by subtle chemical and mineralogical changes to major minerals and the lherzolite bulk composition, and by release of CO_2 -rich fluid.

We have identified unequivocal examples of mantle-derived carbonatite metasomatism in spinel lherzolite xenoliths in

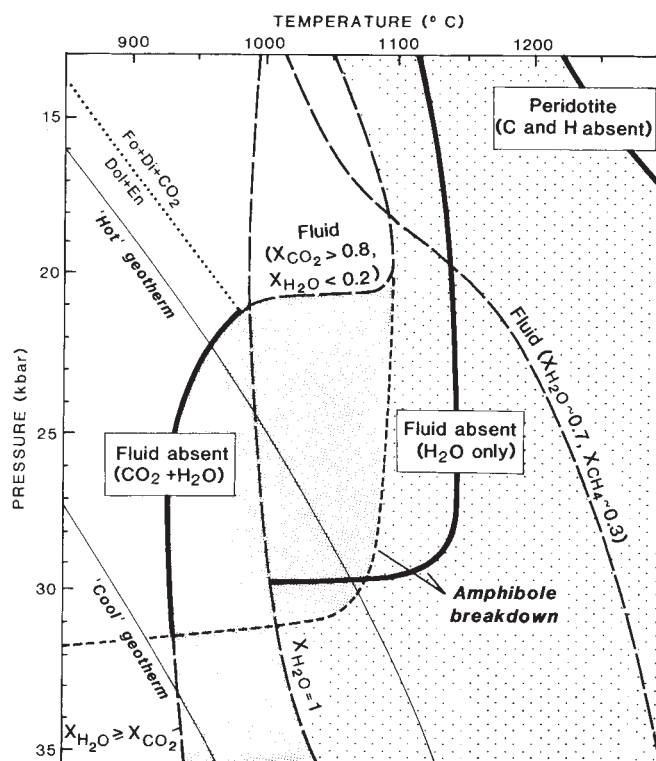


Fig. 1 Solidi for peridotite-C-H-O, illustrating the contrasted effects of reduced fluids ($\text{H}_2\text{O}-\text{CH}_4$)³¹ and oxidized fluids ($\text{H}_2\text{O}-\text{CO}_2$). At $P < 21$ kbar the solidi lie between the peridotite (volatile-free) solidus and the minimum melting temperature for volatile excess ($X_{\text{H}_2\text{O}} = 1$) peridotite- H_2O ¹⁰. But at $P > 21$ kbar, sub-solidus carbonate stability results in a lower temperature solidus for fluid-absent dolomite and amphibole lherzolite¹. The heavily shaded area defines the field of carbonatite melt in equilibrium with amphibole garnet lherzolite ($P < 30$ kbar) or garnet lherzolite ($P > 31$ kbar). At higher temperatures (lightly shaded area) amphibole breaks down and silicate melt is formed, in equilibrium with dolomite-free, amphibole-free lherzolite or garnet lherzolite residue. It is emphasized that these melts are fluid-undersaturated (thick lines) for the amphibole and dolomite lherzolite except at near-solidus conditions at $P < 21$ kbar (CO_2 -rich fluid) or at $P > 31$ kbar, $T < 1,000$ °C. These relationships are valid for fertile lherzolite composition with $\text{H}_2\text{O} < 0.4\%$ and $\text{CO}_2 < 2.5\%$ approximately¹.

basanite from Mt Leura, Victoria^{2,23}. Partial replacement of enstatite by magnesian diopside containing 1.5-2% Na_2O is observed, accompanied by apatite and minor chrome-spinel and an intergranular unidentified phase (Ca, Na + K aluminosilicate). CO_2 -rich fluid inclusions of several generations are abundant. Petrographically, smaller degrees of metasomatism cause an increase in clinopyroxene/orthopyroxene (cpx/opx) ratios, a decrease in MgAl_2O_4 component in spinel and in $\text{MgAl}_2\text{SiO}_6$ solid solution in pyroxene and, most distinctively, the appearance of accessory apatite. Chemically, the reaction of primary carbonatite melt with spinel lherzolite produces an increase in Ca/Al and Na/Ca ratios and an increase in LILE elements without significant decrease in Mg^* ($\text{Mg}^* \equiv 100\text{Mg}/(\text{Mg} + \text{Fe})$ (atomic)). The latter point is important and contrasts with metasomatism by addition of highly undersaturated silicate melt⁴. The absence of a significant shift in Mg^* results from the small Mg/Fe fractionation between residual silicates (Fo_{88}) and alkaline dolomitic carbonatite melt ($\text{Mg}^* = 85$)¹ compared with silicate melt at $\text{Mg}^* = 72$.

The end result of primary carbonatite metasomatism at $P < 21$ kbar is to convert spinel lherzolite to apatite-bearing wehrlite containing magnesian olivine and diopside. Spinel + apatite-bearing wehrlites have been described from mantle-

Table 1 Sub-solidus assemblages crystallized by reacting 2% sodic, dolomitic carbonatite melt¹ with amphibole pyrolite, compared with similar phases in CO₂-free amphibole pyrolite¹⁰

	CO ₂ -free amphibole pyrolite at 15 kbar, 970 °C (ref. 9)			T-2539, 98% amphibole pyrolite + 2% carbonatite ¹ , 18 kbar, 1,000 °C, 4 h (includes olivine Mg* = 88)				T-2535, 98% amphibole pyrolite + 2% carbonatite 22 kbar, 900 °C, 26 h (includes olivine Mg* = 87.6)				
	opx	cpx	amph	opx	cpx	amph	ilm	opx	cpx	amph	ilm	dol
SiO ₂	54.28	52.08	46.04	55.75	52.50	45.01	—	55.19	53.23	44.08	0.72	0.37
TiO ₂	0.37	0.71	1.48	0.30	0.73	2.08	54.18	0.20	0.27	1.61	48.66	—
Al ₂ O ₃	3.91	4.02	11.79	2.87	3.31	12.24	0.25	2.17	2.80	14.63	0.77	0.22
Cr ₂ O ₃	0.77	0.82	1.02	0.39	0.62	0.94	1.67	0.29	—	0.55	1.47	—
FeO	9.47	4.40	5.39	7.55	3.94	5.24	30.53	7.77	5.78	5.14	35.78	2.93
MgO	30.15	17.19	18.45	32.20	17.17	18.37	13.37	33.62	16.42	18.50	12.23	17.89
CaO	1.05	20.52	11.26	0.94	21.48	10.90	—	0.76	20.85	10.08	0.37	30.19
K ₂ O	—	—	0.50	—	—	0.45	—	—	—	0.27	—	—
Na ₂ O	—	0.26	2.07	—	0.25	2.77	—	—	0.65	3.14	—	—
Total	99.91	100.00	98.00	100.00	100.00	98.00	100.00	100.00	100.00	98.00	100.00	51.60
Mg*	85.0	87.4	85.9	88.40	88.60	86.20	43.80	88.5	83.5	86.5	37.9	91.6

amph = amphibole, ilm = ilmenite, dol = dolomite, opx = orthopyroxene, cpx = clinopyroxene.

xenolith suites^{2,25}, but no satisfactory explanation of their enrichment in apatite and Na and Mg-rich clinopyroxenes and of their trace-element characteristics was possible within a conceptual model of upper-mantle metasomatism that included only undersaturated, H₂O-rich and Ti-rich silicate melts, and C-H-O fluids. The characteristics of sodic, dolomitic carbonatite melt and its ephemeral nature in metasomatism (that is, ~50% by weight of the original melt is lost as CO₂-rich fluid, and ~50% enters solid solution in existing stable phases) offer a satisfying explanation of these earlier observations. With reference to the chemical trends from 'primitive' to 'residual' compositions observed in spinel lherzolite suites, the evidence for extensive primary carbonatite metasomatism in the mantle can be seen in distinctive cross-trends to high normative and modal cpx/opx ratios, and high Ca/Al ratios, at Mg* ≥ 89, evident in some data sets²⁶⁻²⁸. It is also important to note that this distinctive metasomatism is characteristic of the spinel-lherzolite, as opposed to garnet-lherzolite, stability field for fertile and depleted mantle lherzolites¹.

The distinctive melting behaviour of carbonated pargasite peridotite¹ is applicable only to mantle redox conditions at $f_{O_2} \geq \text{EMOG}$ (enstatite + magnesite + olivine + graphite) buffer, because at lower oxygen fugacity conditions, CO₂ or carbonate is reduced to graphite. If mantle f_{O_2} conditions are variable²⁹⁻³⁴ then considerable complexity in fluid mobility and partial melting will result (Fig. 1). There is little effect on pargasitic amphibole stability relations for peridotite-H₂O or peridotite-C-H-O over the range IW (iron-wüstite) + 1 log unit < f_{O_2} < EMOG. But the solidus at low f_{O_2} (< EMOG) is very sensitive to f_{O_2} , because decreasing f_{O_2} rapidly changes the fluid or potential fluid phase from H₂O + CO₂, through H₂O > CO₂ + CH₄ to H₂O > CH₄ as $f_{O_2} \rightarrow \text{IW} \pm 1$ log unit ($P = 25-30$ kbar, $T = 900-1,200$ °C). We emphasize that the phase relationships and consequent processes of metasomatism that we have outlined are applicable to relatively oxidized regions of the lithosphere/asthenosphere, that is, $f_{O_2} \geq \text{EMOG}$, and for this reason may be particularly applicable in the lithospheric wedge above Benioff Zones^{32,34}. It is in this region that petrologists and geochemists have sought mechanisms for enrichment in LILE without concomitant enrichment in HFSE: this separation may be achieved by the metasomatic process described herein for the (relatively) oxidized peridotite-C-H-O system.

If it were possible to generalize on a regional scale for the Earth's lithosphere and upper mantle then we would predict that the deeper oceanic lithosphere (>30-40 km?) has low f_{O_2} such that either $f_{CH_4} \approx f_{H_2O}$ or $f_{H_2O} > f_{CH_4} > f_{CO_2}$. Pargasitic amphibole would be stable in lherzolitic compositions (Fig. 1), but sodic, dolomitic carbonatite melts would only occur under locally higher f_{O_2} conditions. Silicate melts (asthenospheric melts) may occur at pressures greater than pargasite stability,

that is, at depths ≥ 90 km (ref. 12), provided that f_{CH_4} is not too high³¹. By contrast, continental lithosphere is envisaged as having grown through geological time, to have suffered rifting and re-suturing, and thus to be underplated by slabs and slivers of old oceanic crust and lithosphere. We predict sub-continental lithosphere to have variable but, on average, higher f_{O_2} (≥ EMOG). Primary carbonatites may be typical of large regions of shallow sub-continental lithosphere (65-95 km), under conditions of moderate to high geothermal gradients.

The distinctive petrographic and chemical characteristics of lithosphere metasomatized by primary sodic, dolomitic carbonatite melts have been recognized in natural spinel lherzolites and wehrlites from continental lithosphere beneath Victoria, Australia, and an examination of literature descriptions suggests their occurrence in many similar xenolith suites. We infer that mantle metasomatism by ephemeral primary carbonatite magmas is an effective process causing local depletions ($P > 20$ kbar) and enrichments ($P < 21$ kbar) in LILE, phosphorus and some HFSE within the lithosphere, particularly beneath continental regions.

We thank K. L. Harris and W. Jablonski for technical assistance; W. R. Taylor, J. Adam and A. J. Stolz for comments on earlier drafts; and J. Meen for helpful comments. Research has been funded by the Australian Research Grants Committee (D.H.G.) and a Commonwealth Post-Graduate Scholarship (M.E.W.).

Received 3 August; accepted 30 October 1988.

- Wallace, M. E. & Green, D. H. *Nature* **335**, 343-345 (1988).
- Frey, F. A. & Green, D. H. *Geochim. cosmochim. Acta* **38**, 1023-1029 (1974).
- Irving, A. J. *Am. J. Sci.* **A** **280**, 389-426 (1980).
- Lloyd, F. E. & Bailey, D. K. in *Physics and Chemistry of the Earth* Vol. 9 (eds Ahrens, L. H., Dawson, J. B., Duncan, A. R. & Erlank, A. J.) 389-416 (Pergamon, London, 1975).
- Menzies, M., Kempton, P. & Dungan, M. *J. Petrol.* **26**, 663-693 (1973).
- McDonough, W. F., McCulloch, M. T. & Sun, S. S. *Geochim. cosmochim. Acta* **49**, 2051-2067 (1985).
- Zindler, A., Staudigel, H. & Batiza, R. *Earth planet. Sci. Lett.* **70**, 175-195 (1984).
- Menzies, M. A. & Murthy, V. R. *Am. J. Sci.* **A280**, 662-688 (1980).
- Menzies, M. A. & Hawkesworth, C. J. (eds) *Mantle Metasomatism* 1-472 (Academic, London, 1987).
- Green, D. H. *Earth planet. Sci. Lett.* **19**, 37-55 (1973).
- Green, D. H. *Tectonophysics* **17**, 456-465 (1973).
- Green, D. H. & Lieberman, R. C. *Tectonophysics* **32** 61-92 (1976).
- Olafsson, M. & Eggler, D. H. *Earth planet. Sci. Lett.* **64**, 305-315 (1983).
- Meen, J. K. *Spec. Pap. geol. Soc. Am.* **215**, 91-100 (1987).
- Meen, J. K., Ayers, J. C. & Fregean, E. J. in *Carbonatites* (ed. Bell, J. K.) (Allen & Unwin, in the press).
- Menzies, M. A., Rogers, N., Tindle, A. & Hawkesworth, C. J. in *Mantle Metasomatism* (eds Menzies, M. A. & Hawkesworth, C. J.) 313-361 (Academic, London, 1987).
- Kempton, P. D. in *Mantle Metasomatism* (eds Menzies, M. A. & Hawkesworth, C. J.) 45-89 (Academic, London, 1987).
- Schneider, M. E. & Eggler, D. H. in *Kimberlites I: Kimberlites and Related Rocks* (ed. Kornprobst, J.) 383-394 (Elsevier, Amsterdam, 1984).
- O'Reilly, S. J. in *Mantle Xenoliths* (ed. Nixon, P. H.) 661-670 (Wiley, London, 1987).
- Ryabchikov, I. D., Schreyer, W. & Abraham, K. *Contr. Miner. Petrol.* **79**, 80-84 (1982).
- Stolz, A. J. *Miner. Mag.* **51**, 719-732 (1987).
- Twyman, J. D. & Gittins, J. in *Alkaline Igneous Rocks* (eds Fitton, J. G. & Upton, B. J.) 85-94 (Blackwell, London, 1987).

23. Deans, T. & Roberts, B. *J. geol. Soc. Lond.* **141**, 563–580 (1984).
24. Dawson, J. B., Garson, M. S. & Roberts, B. *Geology* **15**, 765–768 (1987).
25. Kleeman, J. B., Green, D. H. & Lovering, J. F. *Earth planet. Sci. Lett.* **5**, 449–458 (1969).
26. Nickel, K. G. & Green, D. H. in *Kimberlites II: the Mantle and Mantle–Crust Relationships* (ed. Kornprobst, J.) 161–178 (Elsevier, Amsterdam, 1984).
27. Maaloe, S. & Aoki, K. *Contr. Miner. Petrol.* **63**, 161–173 (1977).
28. O'Reilly, S. Y. & Griffin, W. L. in *Mantle Xenoliths* (ed. Nixon, P. H.) 267–280 (Wiley, London, 1987).
29. Ryabchikov, I. D., Green, D. H., Wall, V. J. & Brey, G. P. *Geochem. Int.* **1981**, 148–158 (1981).
30. Eggler, D. H. *Geophys. Res. Lett.* **10**, 365–368 (1983).
31. Taylor, W. R. & Green, D. H. *Nature* **332**, 349–352 (1988).
32. Green, D. H., Falloon, T. J. & Taylor, W. R. in *Magmatic Processes: Physicochemical Principles* (ed. Mysen, B. O.) 139–154 (Geochem. Soc., London, 1987).
33. Woermann, E. & Rosenhauer, M. *Fortschr. Miner.* **63**, 263–349 (1985).
34. Green, D. H., Taylor, W. R. & Foley, S. in *Stable Isotopes and Fluid Processes in Mineralization* (ed. Herbert, H. K.) (Blackwell, London, in the press).

Is subducted lithosphere trapped at the 670-km discontinuity?

U. Christensen

Max-Planck-Institut für Chemie, Saarstrasse 23, 6500 Mainz, FRG

Recent work^{1,2} on phase equilibria has shown that subducted oceanic lithosphere, which is differentiated into a basaltic crust and a harzburgitic residue layer, could be significantly denser than undifferentiated peridotite at depths between 650 and 700 km, but would be less dense in the range 700–740 km. It might thus be trapped in a gravitationally stable layer between upper and lower mantle³. Here, using numerical experiments of isoviscous convection, I explore the stability of such a layer, assuming that it had previously been formed. The results suggest that the survival of the intermediate layer is at best marginally feasible, provided that new subducted lithosphere is efficiently trapped.

At most depths in the mantle the crustal part of the lithosphere, which is transformed to eclogite, is denser than material of peridotitic composition, whereas the harzburgite is less dense. Both combined are close to being neutrally buoyant, aside from thermal effects. A remarkable exception occurs near the 670-km discontinuity as a consequence of the sensitivity of the garnet-to-perovskite transition to the aluminium content^{1,2}. At about 670 km depth the transformation to perovskite is completed in harzburgite, but not in peridotite, which makes an assemblage of eclogite and harzburgite ~2% denser than peridotite. At 720 km the transition is completed in peridotite, but not in eclogite, which makes the assemblage buoyant at this

depth. Ringwood and Irifune³ infer that at least young subducted lithosphere would be trapped to form a global transition layer between upper and lower mantle. The layer would be occasionally punctured by old slabs, giving rise to a hybrid form of mantle circulation.

Here the mechanism of trapping subducted lithosphere is not considered; rather, the stability of the intermediate layer is studied, assuming that it had been formed. Solutions for isoviscous time-dependent convection, heated from below in two-dimensional boxes of aspect ratio 3 and 2.5, are calculated numerically for thermal Rayleigh numbers 3×10^5 and 10^6 . With such parameters, strongly time-dependent flow occurs with blobs detaching randomly from the boundary layers^{4,5}. Ten thousand tracer particles are used to indicate the distribution of the chemically distinct material⁶, assumed to be an inseparable mixture of eclogite and harzburgite. The tracer particles are neutrally buoyant at most depths z , except in a layer of half-width w centred at z_0 . There the density anomaly $\Delta\rho_c$, caused by a standard concentration of tracers, scaled to the thermal density anomaly $\Delta\rho_T = \rho\alpha\Delta T$, is given by

$$\Delta\rho_c/\Delta\rho_T = R_p 3\sqrt{3}/2 \left\{ \frac{z-z_0}{w} - \left(\frac{z-z_0}{w} \right)^3 \right\},$$

$$\left| \frac{z-z_0}{w} \right| \leq 1$$

where α is the thermal expansion coefficient and ΔT the temperature contrast across the whole convection layer. R_p is the ratio of the maximum chemical buoyancy to the maximum thermal buoyancy. It is about 0.4 in the Earth, assuming values of $\alpha = 2.5 \times 10^{-5} \text{ K}^{-1}$ and $\Delta T = 2,000 \text{ K}$. The model experiments are started from fully developed thermal convection cells. At time zero, the chemical layer with the standard concentration of tracers is inserted into the region where it is most stable, namely the range between $z_0 - w/2$ and $z_0 + w/2$. I take $z_0 = 0.55h$, h being the height of the whole box. The initial lateral range is from $x = 2w$ to $x = L - 2w$, L being the length of the box. Vertical convection currents can pass through the gaps at the sidewalls.

The Rayleigh number Ra of the models is less than the real mantle value of $\sim 10^7$. Accordingly, the boundary layer thickness (or thickness of the lithosphere), which scales as $Ra^{-1/3}$, is increased in the model. To obtain a similar dynamical behaviour at the different Rayleigh numbers, the width of the stability region of the eclogite–harzburgite assemblage must be increased in the same proportion. Values of $w = 0.1h$ and $0.0666h$ were taken for $Ra = 3 \times 10^5$ and 10^6 , respectively, slightly more than the scaled mantle values from ref. 2. Velocities scale approximately with $Ra^{2/3}$. To translate model time to geological time, on the basis that one overturn in the model would correspond to one overturn in the mantle, the model time must be scaled down by a factor of $(Ra_{\text{model}}/Ra_{\text{mantle}})^{2/3}$. Roughly the same result is obtained when extrapolating the data obtained for either value of the Rayleigh number to the mantle value, which supports the applied principles of scaling.

Different values of R_p , including unrealistic ones, were considered in order to study the principal dynamics of such a system. With $R_p = 5$, which is 10 times the mantle value, strict double-layer convection is observed. At $R_p = 2.3$ (Fig. 1), the transition layer still shows a strong tendency to shut the gaps through which the vertical flow passes and also to choke the flow, but this is at the expense of a continuous loss of material which becomes entrained into the remaining current. Some of the blobs that detach from the boundary layers are stopped by the intermediate layer, whereas others punch through, episodically removing large pieces. A situation temporarily resembling a hybrid form of single- and double-layer convection is observed (Fig. 1). Retrapping of material once removed from the transition layer occurs occasionally, but does not seem to be of major

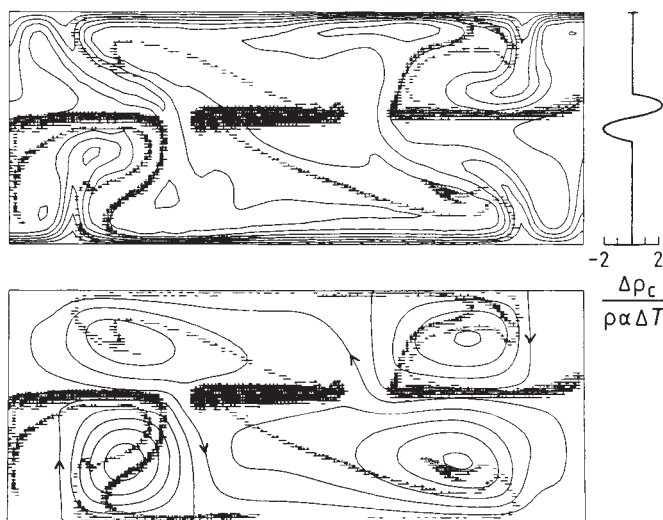
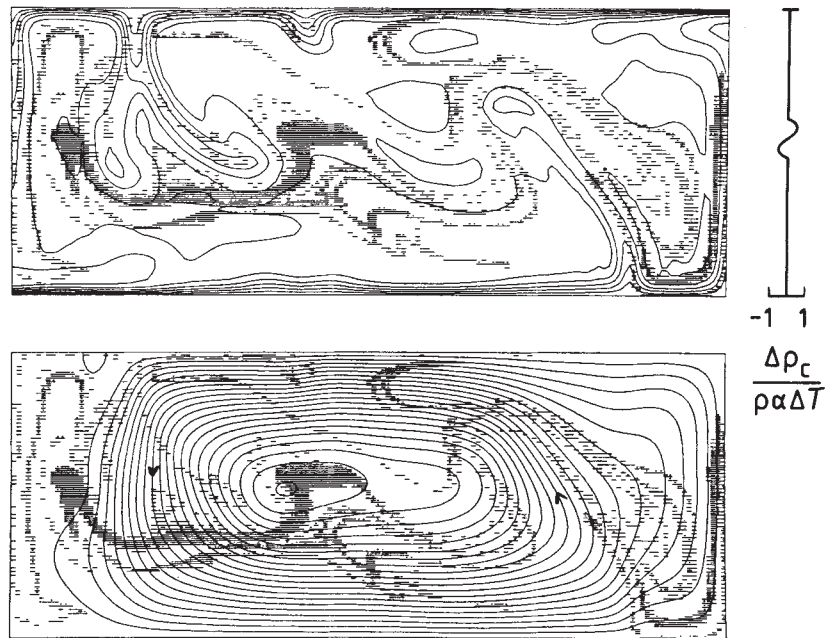


Fig. 1 Isotherms (upper diagram) and stream-lines (lower diagram) in a model with $Ra = 3 \times 10^5$ and $R_p = 2.3$ after about 250 Myr. The distribution of the eclogite–harzburgite assemblage is shaded. Its anomalous density (compared with peridotite) as a function of depth is indicated at the margin.

Fig. 2 As Fig. 1, but for a case of $R_p = 0.5$ and $Ra = 10^6$ after about 250 Myr.



importance for the mass balance. The half-life of anomalous material in the transition layer (without fresh additions) would be of the order of 500 Myr, with $R_p = 2.3$ (Fig. 3). When R_p is decreased to a realistic value of 0.5, the transition layer does not seriously block the mass exchange between upper and lower mantles. Parts of the trapped layer can survive for some time in the stagnating cell core, but are ultimately removed by thermal blobs detaching from the boundary layers (Fig. 2). However, the residence time of the eclogite-harzburgite assemblage in its stability region is increased by a factor of 5–10, as compared with the control case where it would be neutrally buoyant at all depths (Fig. 3). The half-life of the transition layer is now ~ 200 Myr.

Conceivably, a negative Clapeyron slope of the transition to

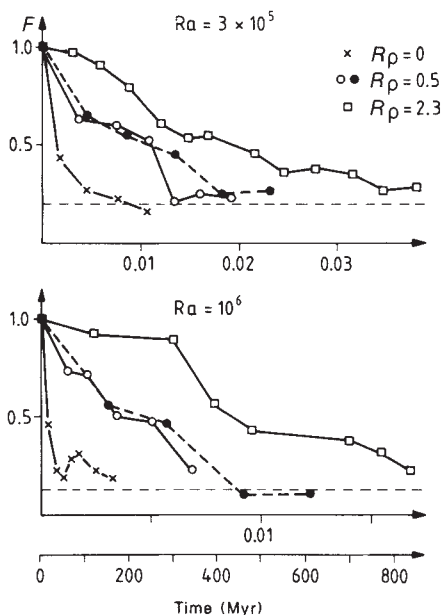


Fig. 3 Fraction (F) of the eclogite-harzburgite assemblage found in its stability region as a function of time. Time is scaled by the thermal diffusion time. The geological timescale at the bottom is appropriate for both values of the Rayleigh number. The horizontal broken lines refer to a random distribution over the whole depth range. Full circles are for a model taking the negative Clapeyron slope into account (no other models do so).

perovskite could further stabilize the intermediate layer. Its influence has been modelled by the analogue of a depth-dependent effective α (ref. 7). This closely reproduces the dynamical influence of a phase transition⁸: when the temperature in the transition region is increased, the phase equilibrium is shifted towards the denser high-pressure phase, which is equivalent to a negative thermal expansion coefficient. A Clapeyron slope of -3 MPa K^{-1} is assumed, which is slightly more negative than experimental values⁹. Although transient blobs are now more effectively impeded from crossing the transition zone, the effect on the residence time is very minor (Fig. 3).

Taking 200 Myr for the half-life of differentiated material in a 60-km-wide transition layer, an annual subduction rate of oceanic lithosphere of 3 km^2 , and 30 km for the thickness of its differentiated part, some 70% of the subducted lithosphere must contribute to replenish the transition layer to balance the losses. Less idealized model calculations that consider variable viscosity are needed to decide whether such efficient trapping is feasible. This is doubtful, because the high viscosity of the slab may not allow its differentiated part to separate from the undifferentiated peridotite, and also it would enable the distributed driving buoyancy of the slab to overcome the localized resisting buoyancy in the transition zone. The present simplified calculations suggest that the survival of a global layer of trapped former lithosphere is at best marginally feasible in a mantle of otherwise uniform composition. In general, this agrees with the conclusion from an earlier model study on the settling of dense crust to the bottom of the convection layer⁶, namely that the intrinsic buoyancy is too small to dominate over thermal buoyancy. But a regional and temporary existence of patches of trapped material cannot be ruled out. Also, if there is a compositional difference between upper and lower mantle which inhibits significant mass flux across the boundary, the gravitational trapping mechanism might work more efficiently.

Received 24 August; accepted 31 October 1988.

1. Irifune, T. & Ringwood, A. E. in *High Pressure Research in Geophysics* (eds Manghnani, M. & Syono, Y.) 235–246 (American Geophysical Union, Washington, 1987).
2. Irifune, T. & Ringwood, A. E. *Earth planet. Sci. Lett.* **86**, 365–376 (1987).
3. Ringwood, A. E. & Irifune, T. *Nature* **331**, 131–136 (1988).
4. Hansen, U. & Ebel, A. *Geophys. J.* **94**, 181–192 (1988).
5. Christensen, U. *Geophys. Res. Lett.* **14**, 220–223 (1987).
6. Gurnis, M. *J. geophys. Res.* **91**, 11407–11419 (1986).
7. Schubert, G., Yuen, D. A. & Turcotte, D. L. *Geophys. J. R. astr. Soc.* **42**, 707–735 (1975).
8. Christensen, U. & Yuen, D. A. *J. geophys. Res.* **90**, 10291–10300 (1985).
9. Ito, E. & Yamada, H. in *High Pressure Research in Geophysics* (eds Akimoto, S. & Manghnani, M.) 405–419 (Center of Academic Publishing, Tokyo, 1982).

Evidence from the Swartkrans cave for the earliest use of fire

C. K. Brain* and A. Sillen†

* Transvaal Museum, PO Box 413, Pretoria 0001, South Africa

† Department of Archaeology, University of Cape Town, Rondebosch 7700, South Africa

During recent excavations of hominid-bearing breccias in the Swartkrans cave altered bones were recovered from Member 3 (about 1.0–1.5 Myr BP) which seemed to have been burnt. We examined the histology and chemistry of these specimens and found that they had been heated to a range of temperatures consistent with that occurring in campfires. The presence of these burnt bones, together with their distribution in the cave, is the earliest direct evidence for use of fire by hominids in the fossil record. Although abundant remains of *Australopithecus robustus* and *Homo cf. erectus* are found in the older Members 1 and 2 at Swartkrans, there is no evidence of fire, suggesting that the discovery of fire was made in the interval between Members 2 and 3 and before *A. robustus* became extinct.

Some distinctive characteristics of burnt bone and tooth surfaces have been previously documented¹, but the surfaces of Swartkrans fossils are often obscured by heavy encrustations of calcium carbonate and manganese dioxide, which make such identification of burnt bones difficult. Therefore, it was necessary to establish what histological and chemical changes appear in fresh bones when experimentally heated so that these effects could then be compared with those in altered Swartkrans fossils.

To provide reference specimens, a fresh hartebeest (*Alcelaphus buselaphus*) radius was cleaned and cut transversely into seven segments. Individual segments were heated to temperatures ranging from 200–800 °C for 30 min, then cooled slowly. Thin sections were prepared and examined by light microscopy. In bones heated to 300–400 °C, there was overall darkening and the lamellar structure was greatly accentuated. Such accentuation was also apparent in blackened bone from Swartkrans Member 3, but not in normal Swartkrans fossils. Bones in the experimental series that were subjected to temperatures higher than 400 °C undergo more complete combustion, revert to a pale colour and show severe cracking with loss of structural detail.

As bones from southern African cave sites may be blackened from diagenetic incorporation of manganese dioxide, we had to establish whether blackening in this case was due to the presence of char. Thus, in addition to histological examination, fossil specimens were characterized by CHN analysis. In a series of 10 specimens, darkened Swartkrans fossils always contained over 2% carbon, whereas normal fossils always contain less than 1.5% carbon (Table 1). A background level of approximately 1% carbon is associated with the inorganic phase; organic char is recoverable from only darkened fossils (Table 1).

We found that, in the experimental series, such char develops mainly between 300 and 400 °C. The carbon-to-nitrogen ratio (C:N) of fresh collagen ranges from 2.9 to 3.6 (ref. 2); for char in the experimental series it is 4.2–6.0. Swartkrans fossil chars are even more depleted in nitrogen; this is consistent with elevated C:N ratios reported for organic residues from other burnt archaeological bones².

Hominid remains from Swartkrans appear in three successive stratigraphic units, Members 1–3. The oldest of the deposits, Member 1, is divided into two discrete masses: the Hanging Remnant, a vast block of fossiliferous breccia clinging unsupported to the cave's north wall, and the Lower Bank resting on the boulder-choked cave floor. On the basis of faunal considerations, both elements of Member 1 are thought to date from 1.8–1.6 Myr BP (ref. 3). Between the Lower Bank and the Hanging Remnant is an erosional space up to 5 m high into which

Table 1 CHN analyses of Swartkrans Member 3 fossils and experimentally heated bones

Normal Member 3 fossil bones					
Whole specimen	% in treated powder				
	C	H	N		
SKX 29281	1.1	0.2	0.1		
SKX 30074	1.2	0.2	—		
SKX 30205	1.2	0.2	—		
SKX 28952	0.7	0.2	—		
SKX 29756	1.0	0.3	—		
Mean	1.0*				
Darkened Member 3 fossil bones					
SKX 26130	2.2	0.3	0.2		
SKX 26131	3.1	0.5	0.3		
SKX 26136	2.8	0.3	0.3		
SKX 26151	2.3	0.3	0.2		
SKX 26126	2.3	0.4	0.2		
Mean	2.5*				
Residues from experimentally heated bone series					
Sample	% weight of residue	% composition of residue			C: N
		C	H	N	
Unheated	1.4	72.0	8.9	trace	
200 °C	0.8	74.0	9.3	trace	
300 °C mean	2.6	58.1	3.8	11.7	5.0
400 °C mean	2.4	58.8	3.8	13.4	4.2
500 °C	1.2	I			
600 °C	0.8	I			
700 °C	0.3	I			
800 °C	0.0	I			
Darkened Swartkrans Member 3 fossil chars					
SKX 26157	1.3	25.7	2.5	3.2	8.0
SKX POOL	1.6	33.8	2.5	2.9	11.7
Normal Swartkrans Member 3 fossils					
SKX 28952	No residue				
SKX 30074	No residue				
SKX 29756	No residue				

Whole specimens were pre-treated in 1M acetic acid to remove normal carbonates. A baseline of approximately 1% carbon is seen in all Swartkrans fossils due to structural apatitic carbonate. After demineralization and hydrolysis, Swartkrans and experimental chars were washed, freeze-dried, weighed and analysed using a Heraeus CHN-Rapid thermal conductivity detector.

* Significant difference using Student's *t* test: $P < 0.001$.

I, insufficient residue for analysis.

heavily calcified sediment (Member 2) has been deposited. The next erosional cycle cut a deep gully through both these deposits along the cave's west wall into which Member 3 was deposited⁴.

Both *A. robustus* and *Homo cf. habilis* are known from Members 1 and 2. Only *A. robustus* remains, from nine individuals, have been recovered from Member 3 (ref. 5), although *Homo* is assumed to have been present. Tools of the Developed Oldowan tradition are found throughout the sequence and no important faunal difference exists between Members 1, 2 and 3, suggesting that they are all within the time range 1.0–1.8 Myr BP (ref. 4). Moreover the environmental conditions prevailing during the accumulation of these Members seem to be similar.

Burnt bones, apparently absent in Members 1 and 2, appear recurrently in Member 3: an average of six 10-cm-thick levels per m² contain burnt bones (Fig. 2). In one 1-m² test pit (W3/S3) burnt fragments were found in 20 separate vertical levels. In W3/S3 there is a scatter of bone pieces, fractured stone, bone tools and remains of *A. robustus* throughout the depth of the profile (Fig. 3). The recurrence of burnt fragments in this and other grid-square profiles suggests that fire was a regular event

Fig. 1 The weight percentage of insoluble char in an experimentally heated bone series. Bone specimens were freezer-milled and decalcified with 15% EDTA. Protein was solubilized and removed by hydrolysis in 6 M HCl at 100° for two days. The insoluble residue, consisting of lipids (below 200 °C) and char, was washed, recovered, lyophilized and weighed. Carbonized char is seen primarily in specimens heated to 300–500 °C, and is responsible for observable blackening.

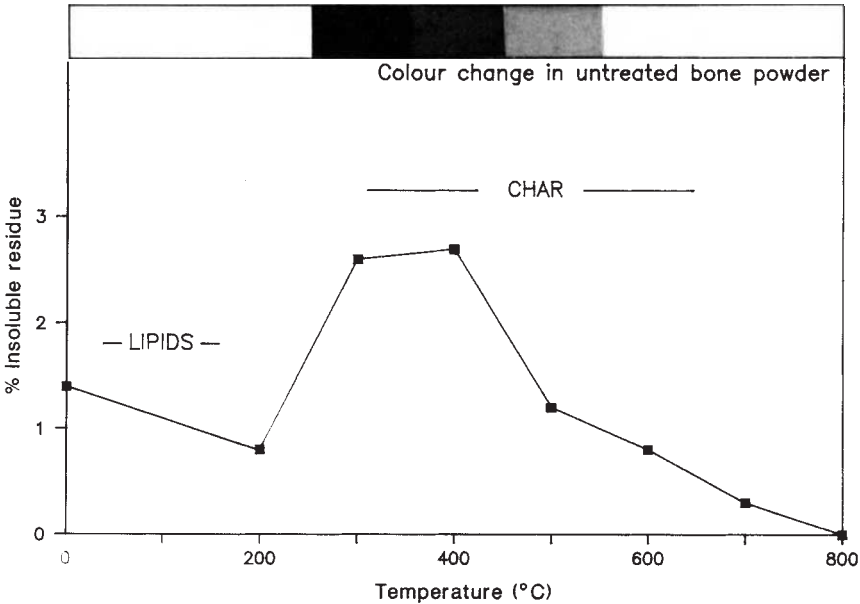


Fig. 2 Plan of the Swartkrans cave showing (a), the position of the Member 3 gully relative to the permanent excavation grid and (b), detail of part of the gully with excavated grid squares hatched. In these squares the upper and lower figures indicate numbers of burnt bone pieces found and total bone pieces respectively.

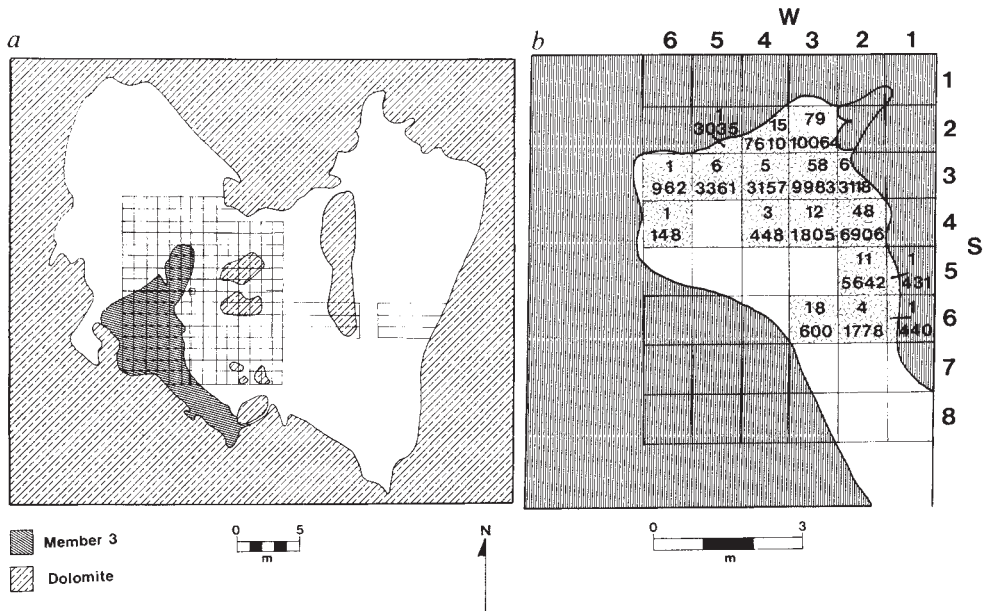
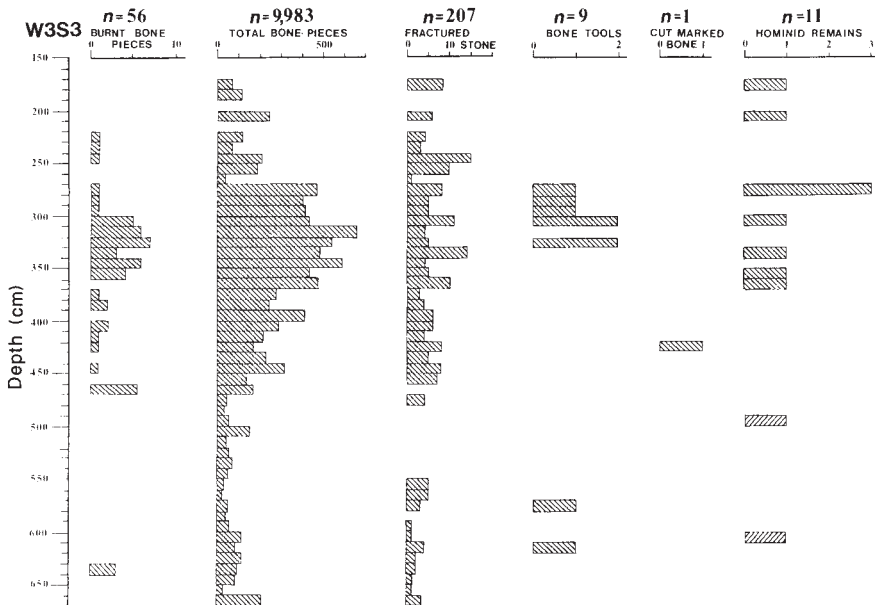


Fig. 3 Profiles through the excavated depth of grid square (W3/S3) in Swartkrans Member 3 sediment. The occurrence of burnt bone pieces relative to total bone pieces, artificially fractured stones, bone tools, cutmarked bone and *A. robustus* remains is indicated.



during the accumulation of Member 3 sediment, rather than an isolated phenomenon.

Excavations of Member 3 yielded a total of 59,488 fossil fragments including 270 pieces judged to have been burnt. Identifiable burnt fragments are mainly derived from antelopes up to the size of wildebeest, however individual burnt specimens of zebra, warthog, baboon and *A. robustus* were also recovered. Of the 270 pieces, it was inferred from colour and structural changes that 46 were lightly heated to below 300°C, 52 to 300–400°C, 45 to 400–500°C and 127 to temperatures above this. These temperatures are consistent with those occurring in experimental campfires made from white stinkwood (*Celtis africana*) branches, the most common tree in the vicinity of Swartkrans.

Indirect evidence for the presence of fire in association with australopithecines exists in the form of magnetic anomalies of sediments at a number of East African localities before 1 Myr BP (refs 6, 7). However, direct evidence for the presence of fire in the form of burnt bones has been lacking. The appearance of burnt bones in Member 3 from Swartkrans is the earliest direct evidence of fire use in the fossil record. Despite similar environmental conditions, no such evidence has yet been found for the period in which Members 1 and 2 were accumulating, suggesting that the discovery of fire occurred in the interval between Members 2 and 3. The presence of cut-marked bones in Member 3 indicate that butchery was practised as well. However the data do not necessarily show that fire was used

for cooking, as protection against predators and the provision of warmth are equally plausible uses. Nor do the data tell us which of the hominids, *Australopithecus*, *Homo*, or both, used the fires. Analysis of suspected burnt specimens from other Plio-Pleistocene localities should help resolve these questions in the future.

We thank Rosemary Newman, Virginia Watson and George Moenda for their help with the Swartkrans investigation; and Julie Lee-Thorp, Portia Johnston and Pierro Benincasa for assistance with the chemical analyses. Many colleagues collaborated on the Swartkrans project, particularly C. S. Churcher, J. D. Clark, E. Delson, F. E. Grine, P. Shipman, R. L. Susman, A. and G. P. Turner. Their help is gratefully acknowledged. The research was supported by grants from the Foundation for Research Development and the University Research Committee of the University of Cape Town. Permission to work at Swartkrans was generously provided by the University of Witwatersrand.

Received 4 July; accepted 3 October 1988.

1. Shipman, P., Foster, G. & Schoeninger, M. *J. Archaeol. Sci.* **11**, 307–325 (1984).
2. De Niro, M. J. *Nature* **317**, 806–809 (1985).
3. Vrba, E. S. in *Hominid Evolution Past Present and Future* (ed. Tobias, P. V.) 195–200 (Alan R. Liss, New York, 1985).
4. Brain, C. K. *et al.* *S. Afr. J. Sci.* (in the press).
5. Grine, F. E. *Proceedings of the International Symposium on the Evolutionary History of Robust Australopithecines* (in the press).
6. Gowlett, J. A., Harris, J., Walton, D. & Wood, B. A. *Nature* **294**, 125–129 (1981).
7. Barbetti, M. J. *hum. Evol.* **15**, 771–782 (1986).

Explosive speciation at the base of the adaptive radiation of Miocene grazing horses

Bruce J. MacFadden & Richard C. Hulbert Jr

Florida Museum of Natural History, University of Florida, Gainesville, Florida 32611, USA

Adaptive radiations are generally considered to be periods of rapid evolutionary change¹. Simpson² viewed rates of evolution as either taxonomic or morphological and various methods for their quantification are widely used^{1,3–5}. In contrast to classic studies of extant organisms^{6–7}, adaptive radiations of fossil groups are usually analysed on the superspecific (for example, generic or familial) level. But relatively few studies have dealt with the patterns and processes of adaptive radiations for fossil species. The principal adaptive radiation of grazing horses occurred in North America 15 to 18 million years ago (Myr BP), when at least 19 species originated by cladogenesis. We report here that the pattern of horse taxonomic evolution is logistic, with very high origination rates observed early during the radiation. By 15 Myr BP diversity reached a plateau of about 16 contemporaneous grazing species, and this taxonomic carrying capacity continued until about 6 Myr BP. In contrast, rates of morphological (dental) evolution during the period were 'normal' (or holotelic) relative to other known fossil groups. Thus, adaptive radiations are times of rapid cladogenesis, but not necessarily rapid morphological change.

Because of their widespread occurrence in space and time, the 55 Myr phylogeny of horses (Equidae) is a familiar example of many evolutionary processes (Fig. 1). It is generally agreed that the transition from moderate to high-crowned dentitions in horses occurred within the advanced species of *Parahippus* and primitive *Merychippus*^{2,8,9}, yet phylogenetic relationships during this important morphological and adaptive transformation are poorly known. We analysed 97 cranial, dental and postcranial characters in 13 critical species both just before (representing the parahippine outgroup) and during the base of the adaptive

radiation of fully hypsodont horses¹⁰. Of these, 40 characters contained phylogenetic information relevant to the present study. Alpha-level species discrimination was based on analysis of both quantitative (usually size) and qualitative morphological characters^{11–13}. The resulting character state/OTU (operational taxonomic unit) matrix was analysed using the computer-based program PAUP¹⁴ to generate the most parsimonious cladogram to express phylogenetic interrelationships. To calculate species ranges and rates of morphological evolution, a recent, precisely calibrated biochronology of North American terrestrial Miocene mammalian faunas was used¹⁵.

Our phylogenetic analysis resulted in a single most parsimonious cladogram involving 146 steps with a consistency index of 0.58. There were 11 cladograms generated at 147 steps and the principal pattern of the cladogenetic events is essentially the same for the first 100 cladograms (up to 149 steps) generated from the PAUP algorithm. The morphological transformation from moderate to high-crowned horses occurred between 18 to 17 Myr BP and is represented by the species sequence '*Parahippus*' *leonensis* (outgroup sister-species) to '*Merychippus*' *gunteri* to '*M.*' *primus* (1, 2 and 3 in Fig. 2). The two tribes of hypsodont horses (Hipparionini and Equini) are each represented by several clades by 17 Myr BP (Fig. 2).

Rates at which species originate are commonly used as measures of taxonomic evolution^{1,2}. Change in the species richness of a clade is a function of its speciation rate (*S*) and extinction rate (*E*), such that the rate of net change (*R*) is the difference between *S* and *E*^{1,3}. This is based on the exponential growth curve, expressed as $R = (\ln N - \ln N_0) / \Delta t$; where N_0 is the initial number of species in a clade and N is the subsequent number of species after elapsed time Δt . For the Equinae, the maximum rate of net growth ($R = 0.97$) occurred during the interval from about 18 to 17 Myr BP (Fig. 3). The value for *R* during the early phase of the radiation is four to five times higher than those of most mammalian clades reported by Stanley¹ (mean value for seven families was 0.22). The standing diversity of hypsodont equids in North America reached a value in the range of 16 ± 3 species by at least 13 Myr BP. It remained in this range for the next 7 Myr, as extinctions balanced originations, resulting in a steady-state diversity pattern¹⁶ (Fig. 3). If these data take into account 'tardy' first appearance datum

Fig. 1 Higher-level (generic) phylogeny of the Equidae indicating the position of the major adaptive radiation of high-crowned horses discussed in text. Adapted from ref. 5.

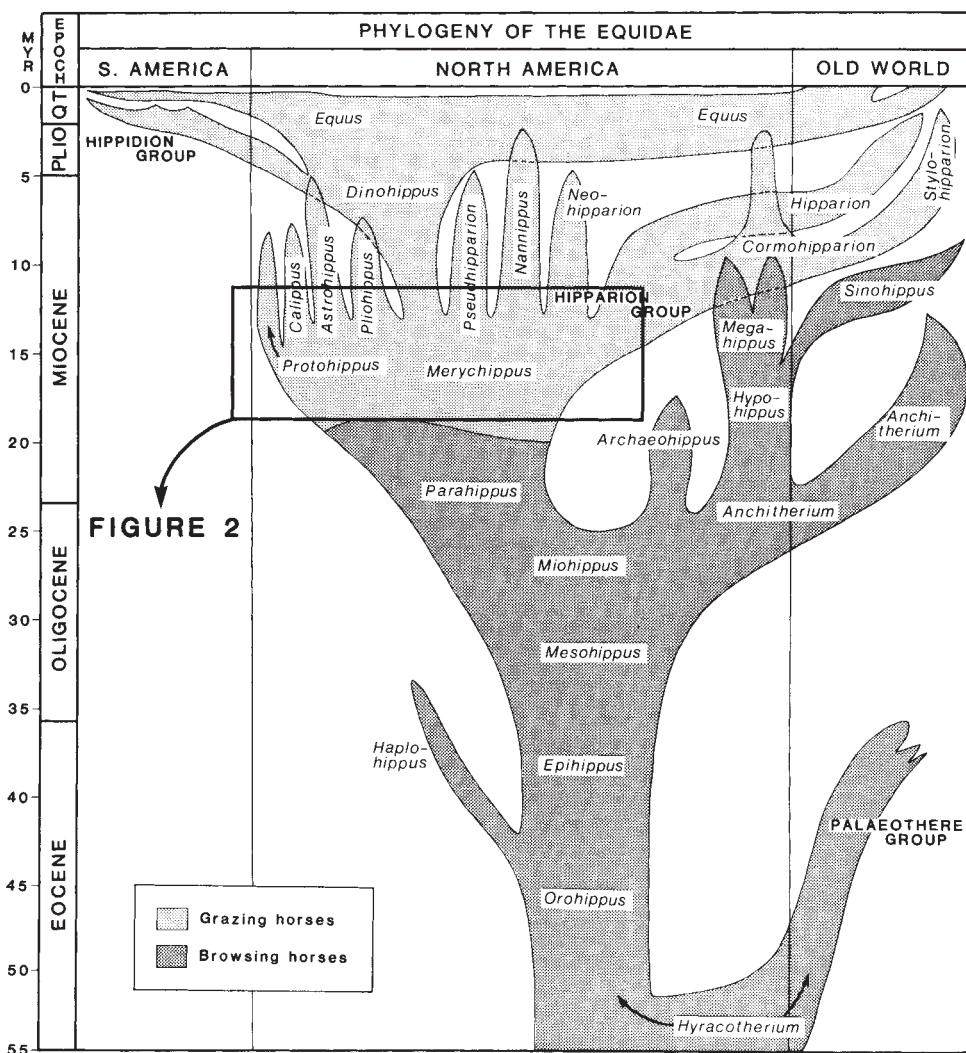
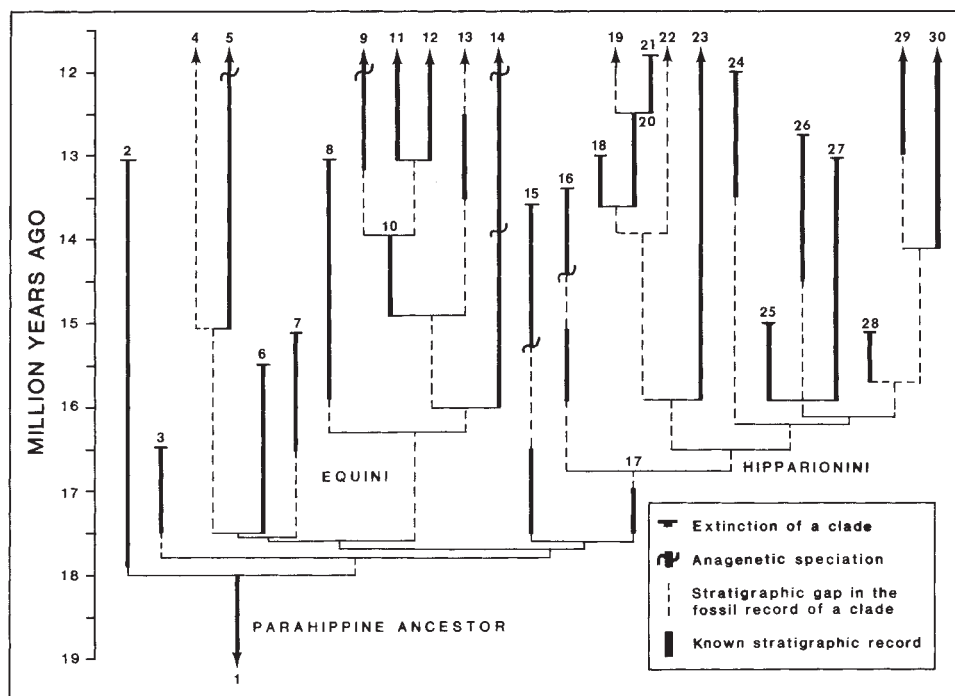


Fig. 2 Phylogenetic tree based on the results of cladistic analysis of horse species at base of adaptive radiation. The numbered clades correspond to the following; 1, 'Parahippus' leonensis; 2, 'Merychippus' gunteri; 3, 'M.' primus; 4, Astrohippus-Dinohippus clade; 5, Pliohippus mirabilis-P. pernix; 6, 'M.' carizzoensis; 7, 'M.' stylodontus; 8, 'M.' intermontanus; 9, Calippus proplacidus-C. placidus; 10, Calippus sp. (Burkeville Fauna); 11, Calippus sp. (lower Bone Valley Formation); 12, Calippus regulus; 13, Calippus circulus; 14, Protohippus vetus-P. perditus-P. supremus; 15, 'M.' tertius-'M.' isonesus; 16, 'M.' sp. (Point Blank Fauna)-'M.' sejunctus; 17, 'M.' sp. (large species, Box Butte Formation); 18, 'M.' republicanus; 19, Pseudhipparion curtivallum and its sister-taxa; 20, Pseudhipparion sp. (lower Bone Valley Formation); 21, Pseudhipparion retrusum; 22, Neohipparion; 23, 'M.' coloradense; 24, Hipparion shirleyi; 25, Merychippus insignis; 26, M. californicus; 27, M. brevidontus; 28, 'M.' goorisi; 29, Nannhippus; 30, Cormohipparion sphenododus. The primary form of the tree is based on the most parsimonious cladogram produced by the PAUP computer program¹⁴ on taxa 2, 3, 5, 6, 8, 14, 15, 16, 22, 23, 24 and 27 using taxon 1 as an outgroup¹⁰. The other taxa were added to this tree based on separate analyses (using PAUP) of major Neogene clades^{12,13,21}.



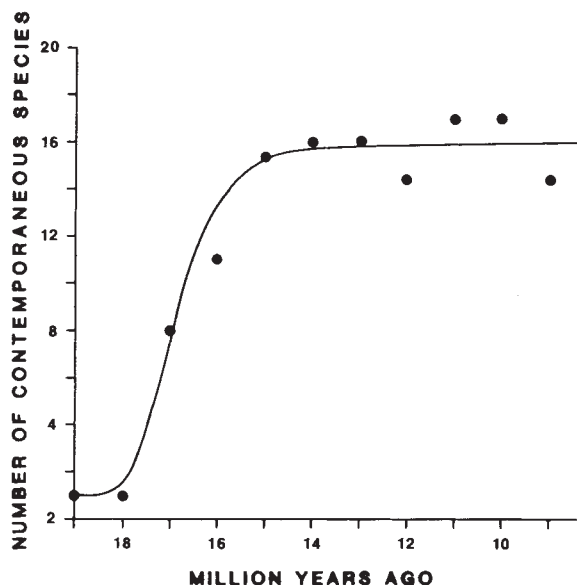


Fig. 3 Plot of species richness/origination for Miocene grazing horses.

planes¹⁷, then this plateau was actually reached about 2 Myr earlier (~15 Myr BP), and R is only 0.06 for the period from 15 to 13 Myr BP. Before 13 Myr BP, speciation was primarily by cladogenesis¹⁰; afterwards, anagenesis became equally important⁵. These results therefore imply an episode of particularly rapid speciation for horses relative to other fossil groups¹.

Haldane¹⁸ introduced the unit 'darwin' (d) to express the rate of evolutionary change of morphological characters as observed in ancestral-descendant fossil taxa. A darwin is calculated by the equation; rate (d) = $(\ln x_2 - \ln x_1)/\Delta t$; where x_1 is the initial size of a character in an ancestral species, x_2 is the subsequent size of the same character in a descendant species; and Δt is the amount of time separating these observations. We have calculated rates of morphological change for the occlusal surface (length and widths) and crown heights of ancestral-descendant species-pairs of fossil horses^{5,19,20}. The rates of dental evolution for the *Parahippus*-*Merychippus* transition at 18–16 Myr BP are between 0.04 to 0.20 d (ref. 20). Although high for horses, these rates are not the highest^{19,20}. Also, they are an order of magnitude lower than those calculated for the early Tertiary mammalian radiation in North America⁴.

In summary, the principal adaptive radiation of hypsodont horses during the Miocene was characterized by high rates of taxonomic evolution (as represented by species origination), but only by moderate (or 'horotelic', *sensu* Simpson²) rates of morphological evolution (as represented by dentitions). If the evolutionary pattern observed for Miocene horses is applicable to other monophyletic groups, then adaptive radiations are times of rapid cladogenetic speciation, but not necessarily rapid morphological change.

We thank Richard H. Tedford for access to the Frick horse collection at the American Museum of Natural History and Douglas S. Jones and S. David Webb for helpful comments on the manuscript. This work was supported by the NSF.

Received 29 June; accepted 14 October 1988.

1. Stanley, S. M. *Macroevolution* (Freeman, San Francisco, 1979).
2. Simpson, G. G. *The Major Features of Evolution* (Columbia University Press, New York, 1953).
3. Stanley, S. M. *Paleobiology* 10, 13–26 (1985).
4. Gingerich, P. D. *Proc. 3rd N. Am. Paleont. Conv.* 1, 205–210 (1982).
5. MacFadden, B. J. *Paleobiology* 11, 245–257 (1985).
6. Lack, D. *Darwin's Finches* (Cambridge Univ. Press, Cambridge, 1947).
7. Greenwood, P. H. *Bull. Br. Mus. (Nat. Hist.), Zool. Suppl.* 6, 1–134 (1974).
8. Matthew, W. D. *Q. Rev. Biol.* 1, 139–185 (1926).

9. Sturton, R. A. *Univ. California Pub. Bull. Dept. Geol. Sci.* 25, 165–198 (1940).
10. Hulbert, R. C. & MacFadden, B. J. *Am. Mus. Novit.* (in the press).
11. MacFadden, B. J. *Bull. Am. Mus. Nat. Hist.* 179, 1–196 (1984).
12. Hulbert, R. C. thesis, Univ. Florida (1987).
13. Hulbert, R. C. *Bull. Florida State Mus., Biol. Sci.* 32, 221–340 (1988).
14. Swofford, D. L. *PAUP, Phylogenetic Analysis Using Parsimony* (Illinois nat. Hist. Survey, Champaign, 1985).
15. Tedford, R. H. in *Cenozoic Mammals of North America* (ed. Woodburne, M. O.) (Univ. California Press, Berkeley, 1988).
16. Cracraft, J. *Syst. Zool.* 31, 348–365 (1982).
17. Novacek, M. J. & Norell, M. A. *Syst. Zool.* 31, 366–375 (1982).
18. Haldane, J. B. S. *Evolution* 3, 51–56 (1949).
19. Webb, S. D. & Hulbert, R. C. in *Vertebrates Phylogeny and Philosophy* (eds Flanagan, K. M. & Lillegraven, J. A.) (Univ. Wyoming, Laramie, 1986).
20. MacFadden, B. J. *Biol. J. Linn. Soc.* 35, 37–48 (1988).
21. Hulbert, R. C. *Bull. Florida State Mus., Biol. Sci.* 33, 229–238 (1988).

Modification of retinal ganglion cell axon morphology by prenatal infusion of tetrodotoxin

David W. Sretavan*‡, Carla J. Shatz* & Michael P. Stryker†

* Department of Neurobiology, Stanford University School of Medicine, Stanford, California 94305, USA

† Department of Physiology, University of California, San Francisco, California 94143-0444, USA

The cellular mechanisms by which the axons of individual neurons achieve their precise terminal branching patterns are poorly understood. In the visual system of adult cats, retinal ganglion cell axons from each eye form narrow cylindrical terminal arborizations restricted to alternate non-overlapping layers within the lateral geniculate nucleus (LGN)^{1–3}. During prenatal development, axon arborizations from the two eyes are initially simple in shape and are intermixed with each other; they then gradually segregate to form complex adult-like arborizations in separate eye-specific layers by birth^{4,5}. Here we report that ganglion cell axons exposed to tetrodotoxin (TTX) to block neuronal activity during fetal life fail to form the normal pattern of terminal arborization. Individual TTX-treated axon arborizations are not stunted in their growth, but instead produce abnormally widespread terminal arborizations which extend across the equivalent of approximately two eye-specific layers. These observations suggest that during fetal development of the central nervous system, the formation of morphologically appropriate and correctly located axon terminal arborizations within targets is brought about by an activity-dependent process.

TTX, which abolishes action potential propagation by blocking voltage-sensitive sodium channels, was continuously infused into the region above the optic chiasm in four fetal cats from embryonic day 42 (E42) to E58 (birth = E65). It is during this period that the eye-specific layers of the LGN normally form as individual retinal axons elaborate a dense terminal arborization in the appropriate layers and lose short side branches in territory destined to belong exclusively to axon arborizations from the opposite eye^{4,5}. In the present experiment, caesarian sections were performed on timed pregnant cats at E42 using inhalation anesthesia and sterile technique⁴. Following surgical exposure, an osmotic minipump (Alza, model 2002, 0.5 $\mu\text{L h}^{-1}$) containing TTX (0.3 mM) in citrate buffer (1.5 mM, pH 4.8) was attached to the back of each fetus. A silastic catheter extending from the outflow port of the pump was then inserted through the frontal bone just lateral to the sagittal sinus. Following pump implantation, fetuses were returned to the uterus and allowed to undergo further development for 16 days, the limit of reliable minipump function. This mode of delivery results in a cerebro-

‡ To whom reprint requests should be sent. Present address: Laboratory of Neurobiology, Rockefeller University, 1230 York Avenue, New York City, New York 10021, USA.

Fig. 1 Retinogeniculate axon terminal arborizations at E58 in normal fetuses and in fetuses following continuous intracranial infusion of TTX between E42 and E58. *a*, Examples from a normal fetus. The diagram in the middle of the figure shows the position of each axon arborization within the LGN (shown in horizontal section). The axon on the left terminates within layer A and is from the contralateral eye. The axon on the right terminates in layer A1 and is from the ipsilateral eye. Note that each axon originates from the optic tract and runs in a straight trajectory in the posterior to anterior direction within the LGN before elaborating a terminal arborization that is restricted to the appropriate eye-specific layer. Abbreviations: A, layer A; A1, layer A1; C-cx, the C-complex; OT, optic tract. *b*, Examples from TTX treated fetuses. Both axons shown here elaborate abnormally expanded terminal arborizations that are no longer restricted in their greatest extent to the thickness of a normal eye-specific layer. Instead, these TTX-treated axons appear to give off branches throughout much of their traverse across the LGN. In addition, TTX-treated axon arborizations are much wider in their mediolateral extent than axons in normal fetuses.

Methods. Retinal ganglion cell terminal arborizations were labelled by placing the fetal diencephalon including the optic tract and the LGN into a modified slice chamber and bulk-filling optic tract axons with horseradish peroxidase (HRP)⁵. Following incubation for 3–4 h, the tissue was fixed, reacted for HRP histochemistry and individual retinogeniculate axon arborizations were reconstructed from serial 100- μ m thick horizontal sections. To correlate the morphology of individual axons with the global pattern of retinal projection in the same fetus, each animal also received an intraocular injection of tritiated leucine into one eye at E57, 24 h before *in vitro* axon labelling. Then, following analysis of arborization morphology, tritium-sensitive film was applied closely against each section to reveal the autoradiographic labelling pattern from the injected eye⁵.

Results. Retinogeniculate axon terminal arborizations at E58 in normal fetuses and in fetuses following continuous intracranial infusion of TTX between E42 and E58 are shown in Fig. 1. The diagram in the middle of the figure shows the position of each axon arborization within the LGN (shown in horizontal section). The axon on the left terminates within layer A and is from the contralateral eye. The axon on the right terminates in layer A1 and is from the ipsilateral eye. Note that each axon originates from the optic tract and runs in a straight trajectory in the posterior to anterior direction within the LGN before elaborating a terminal arborization that is restricted to the appropriate eye-specific layer. Abbreviations: A, layer A; A1, layer A1; C-cx, the C-complex; OT, optic tract. *b*, Examples from TTX treated fetuses. Both axons shown here elaborate abnormally expanded terminal arborizations that are no longer restricted in their greatest extent to the thickness of a normal eye-specific layer. Instead, these TTX-treated axons appear to give off branches throughout much of their traverse across the LGN. In addition, TTX-treated axon arborizations are much wider in their mediolateral extent than axons in normal fetuses.

Methods. Retinal ganglion cell terminal arborizations were labelled by placing the fetal diencephalon including the optic tract and the LGN into a modified slice chamber and bulk-filling optic tract axons with horseradish peroxidase (HRP)⁵. Following incubation for 3–4 h, the tissue was fixed, reacted for HRP histochemistry and individual retinogeniculate axon arborizations were reconstructed from serial 100- μ m thick horizontal sections. To correlate the morphology of individual axons with the global pattern of retinal projection in the same fetus, each animal also received an intraocular injection of tritiated leucine into one eye at E57, 24 h before *in vitro* axon labelling. Then, following analysis of arborization morphology, tritium-sensitive film was applied closely against each section to reveal the autoradiographic labelling pattern from the injected eye⁵.

spinal fluid TTX concentration of about $1 \mu\text{M}$ ⁶; a level sufficient to block action potentials in the optic nerve of newborn rats⁷ and postnatal cats⁸. A previous study has shown that TTX infused in this way prevents the formation of eye-specific layers in the LGN, whereas citrate buffer alone has no effect⁶.

Experimental fetuses were delivered at E58. To determine the effects of TTX on arborization, retinal axons were filled with horseradish peroxidase (HRP) using an *in vitro* technique⁵. As illustrated in Fig. 1*a*, normal axons at E58 run within the LGN for some distance without branching before giving off relatively narrow terminal arborizations restricted in their longest extent to the thickness of an eye-specific layer⁵. At this age, layer A receives terminal arborizations of axons originating almost exclusively from the contralateral eye, while layer A1 receives those from the ipsilateral eye⁴.

In contrast, all retinogeniculate terminal arborizations developing in the presence of TTX were highly abnormal in morphology. Two from our sample of 28 fully reconstructed axons are shown in Fig. 1*b*. TTX-treated axons have expanded terminal arborizations which branch extensively within a region corresponding to the thickness of several layers, making it impossible to distinguish axons originating from the contralateral and ipsilateral eyes. Correspondingly, intraocular injections of anatomical tracers in TTX-treated fetuses revealed that retinal projections from the two eyes overlap with each other almost completely (see also ref. 6).

The arborizations of TTX-treated axons do not resemble those of normal retinal axons at any stage of fetal development. In particular, these axons are quite unlike normal axons at E42 (the start of TTX infusion), which are very simple and have only a single main branch with a few short side branches⁵. Moreover, TTX-treated axons at E58 have much wider terminal arborizations ($253 \mu\text{m} \pm 88 \mu\text{m}$ (s.d.), $n = 14$) compared to normal axons at the same age ($90 \mu\text{m} \pm 33 \mu\text{m}$ (s.d.), $n = 28$); and a greater total linear extent of axon within the terminal arborization (TTX: $3,910 \mu\text{m} \pm 1,116 \mu\text{m}$ (s.d.), $n = 14$; Normal: $2,520 \mu\text{m} \pm 810 \mu\text{m}$ (s.d.), $n = 28$). Therefore, TTX treatment has not simply frozen axon development at an immature state. Instead, TTX-treated axons grow in an unrestricted way, resulting in abnormally large terminal arborizations.

A distinct group of axons with highly unusual trajectories and branching patterns within the LGN was also observed following TTX treatment. An example of one such axon is shown in Fig. 2. Unlike axons in normal fetuses and the vast majority of axons in TTX-treated fetuses, this axon originates from the optic tract along the lateral aspect of the LGN and runs across

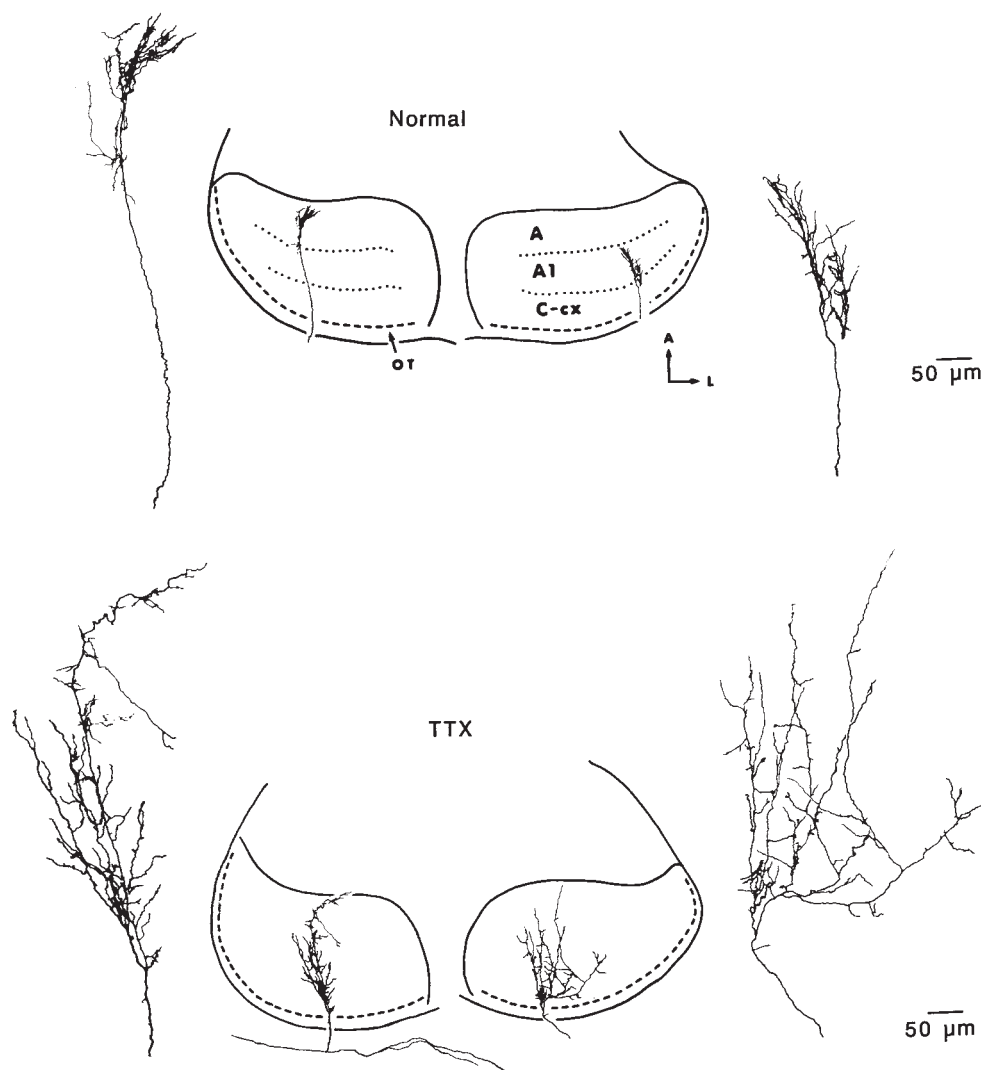
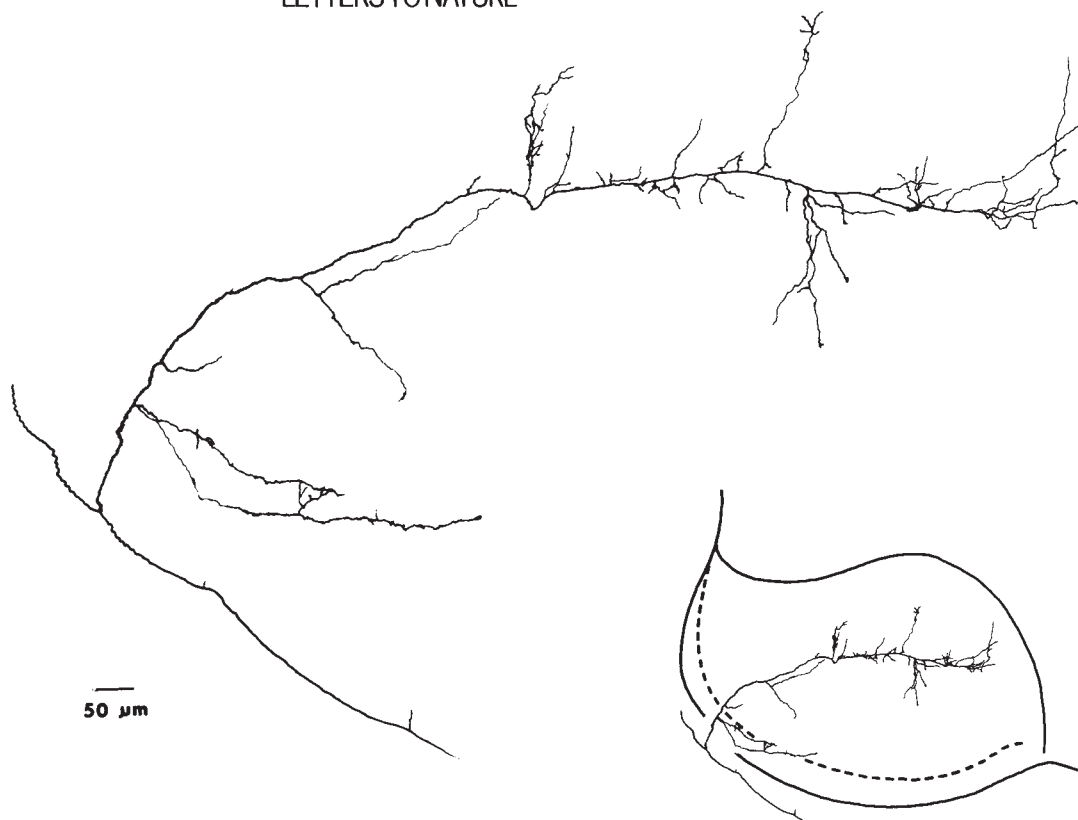


Fig. 2 Example of a TTX-treated axon at E58 with an unusual trajectory and branching pattern. The position of this axon with respect to the LGN is shown in the bottom right corner. In contrast to retinogeniculate axons in normal fetuses and the vast majority of axons in TTX-treated fetuses, this axon travels in a mediolateral direction within the LGN, giving off branches throughout the medial half of the nucleus. Axons with this branching pattern within the LGN have not been seen in normal fetuses²⁰.



the entire nucleus in a mediolateral direction. Numerous branches are given off along its course throughout the medial half of the nucleus. After leaving the LGN, this axon heads posteriorly towards the pretectum and superior colliculus (data not shown). The unusual trajectory taken by this axon within the LGN is perpendicular to that taken by normal fetal retinogeniculate axons (see Fig. 1*a, b* and ref. 5) and would tend to disrupt the normal topography of the retinal map⁹.

These observations indicate that TTX treatment between E42 and E58 affects the development of two basic features of retinogeniculate axon morphology: the restriction of terminal arborizations to appropriate eye-specific layers, and the formation of narrow cylindrical arborizations that underlie the topographic precision of the projection. Studies in many species have implicated neuronal activity in the refinement of connections (see ref. 10 for review). In mammals, physiological studies demonstrate that TTX blocks the formation of ocular dominance columns in layer 4 of the cat's visual cortex⁸ and the segregation of inputs from functionally distinct classes of retinal ganglion cells onto LGN neurons¹¹. Thus far, the effects of activity blockade on the development of individual axon arborizations have only been examined in experimentally created retinal projections in lower vertebrates¹². Results from the present study demonstrate directly that the morphology of axon arborizations during normal prenatal mammalian development is dramatically altered following exposure to TTX. (Preliminary results show that the postnatal development of retinogeniculate arborizations can also be modestly affected by TTX treatment; M. Sur, P. Garraghty and M.P. Stryker; personal communication.)

It is unlikely that TTX affects the development of neuronal connections by gross disruption of axoplasmic transport¹³ because, as shown here, individual ganglion cell axons exposed to TTX undergo extensive growth, exceeding that seen during the same period in normal development by about 50%. Furthermore, axonal transport in TTX-treated fetuses as demonstrated by robust anterograde transport of tracers injected intraocularly is qualitatively normal. It is more likely that TTX exerts its

effects by blocking neuronal activity. The expansion of retinal axon arborizations in the presence of TTX as observed here is consistent with the demonstration that neuronal activity has an inhibitory effect on neurite extension in culture¹⁴ and with recent results showing that activity blockade using neurotransmitter antagonists promotes neurite outgrowth in rat retinal ganglion cells¹⁵.

The alterations in the morphology of arborizations seen in the presence of TTX during fetal cat development are similar to those seen in amphibia. When an additional eye is forced to innervate the frog optic tectum, inputs from the two eyes segregate into stripes reminiscent of mammalian cortical ocular dominance columns¹⁶. Application of TTX intraocularly or to the optic nerves causes desegregation which is manifested at the level of individual axons by an enlargement in branching¹². In lower vertebrates, there is ample evidence to show that TTX blocks visually-evoked action potential activity (for review, see ref. 10), implying that formation of specific patterns of axon arborization requires neuronal activity. In the fetal cat, however, visually-driven activity is absent and photoreceptor outer segments do not mature until about a week postnatally¹⁷. Nevertheless, the machinery necessary for nerve impulse conduction and synaptic transmission is present and even capable of function as early as E38¹⁸. These observations suggest spontaneous action potential activity in the fetus normally drives the process of segregation in the LGN. This suggestion seems even more likely in view of a report by Galli and Maffei¹⁹, who have recorded spontaneous activity *in vivo* from the optic nerve of fetal rats at stages comparable to those studied here.

The normal development of mammalian ganglion cell terminal arborizations within the LGN seems to be a remarkably economical process because only modest elimination of inappropriate branches accompanies the extensive and directed growth of the terminal arborization within the appropriate location⁵. Such economy might be taken to indicate that axonal development in this system is rather rigidly prespecified and is not susceptible to extrinsic, competitive interactions. However, the results repor-

ted here show clearly that the development of arborizations can be perturbed by the intracranial infusion of TTX. Although extensive growth still occurs, the stereotyped and restricted pattern of terminal arborization fails to form. Thus, the highly directed growth seen in the prenatal development of retinogeniculate axon arborizations is likely to arise as a consequence of intense activity-mediated interactions within the LGN. We suggest that axons may use similar activity-dependent interactions to achieve their precise patterns of terminal arborization elsewhere in the developing mammalian CNS.

We thank M. W. Siegel for assistance with fetal surgery. This work was supported by grants from the NSF, the NIH, and the March of Dimes (to C.J.S. and M.P.S.).

Received 1 August; accepted 18 October 1988.

- Hickey, T. L. & Guillery, R. W. *J. comp. Neurol.* **156**, 239–254 (1974).
- Sur, M. & Sherman, S. M. *Science* **218**, 389–391 (1982).
- Bowling, D. B. & Michael, C. R. *J. Neurosci.* **4**, 198–216 (1984).
- Shatz, C. J. *J. Neurosci.* **3**, 482–499 (1983).
- Sretavan, D. W. & Shatz, C. J. *J. Neurosci.* **6**, 234–251 (1986).
- Shatz, C. J. & Stryker, M. P. *Science* **242**, 87–89 (1988).
- Foster, R. E., Connors, B. E. & Waxman, S. G. *Dev. Brain Res.* **3**, 371–386 (1982).
- Stryker, M. P. & Harris, W. J. *J. Neurosci.* **6**, 2117–2133 (1986).
- Sanderson, K. J. *J. Comp. Neurol.* **143**, 101–118 (1971).
- Schmidt, J. T. & Tieman, S. B. *Cell. Molec. Neurobiol.* **5**, 5–35 (1985).
- Dubin, M. W., Stark, L. A. & Archer, S. M. *J. Neurosci.* **6**, 1021–1036 (1986).
- Reh, T. A. & Constantine-Paton, M. *J. Neurosci.* **5**, 1132–1143 (1985).
- Ricco, R. V. & Matthews, M. A. *Neuroscience* **16**, 1027–1039 (1985).
- Cohan, C. S. & Kater, S. B. *Science* **232**, 1638–1640 (1986).
- Lipton, S. A., Frosch, M. P., Phillips, M. D., Tauck, D. L. & Aizenman, E. *Science* **239**, 1293–1296 (1988).
- Law, M. I. & Constantine-Paton, M. *J. Neurosci.* **1**, 741–759 (1981).
- Donovan, A. J. *Comp. Neurol. exp. Eye Res.* **5**, 249–254 (1966).
- Shatz, C. J. & Kirkwood, P. A. *J. Neurosci.* **5**, 1378–1397 (1984).
- Galli, L. & Maffei, L. *Science* **242**, 90–91 (1988).
- Sretavan, D. W. & Shatz, C. J. *J. Comp. Neurol.* **255**, 386–400 (1987).

Positive selection of CD4⁺ thymocytes controlled by MHC class II gene products

H. Robson MacDonald*, Rosemary K. Lees*,
Reto Schneider†, Rolf M. Zinkernagel†
& Hans Hengartner†

* Ludwig Institute for Cancer Research, Lausanne Branch,
1066 Epalinges, Switzerland

† Institute of Pathology, University Hospital,
8091 Zurich, Switzerland

The mature T-cell antigen receptor repertoire is characterized by lack of reactivity to self-components as well as by preferential reactivity to foreign antigens in the context of polymorphic self-proteins encoded within the major histocompatibility complex¹. Whereas the former characteristic (referred to as negative selection or tolerance) is associated with intrathymic deletion of T cells expressing T-cell antigen receptor β -chain variable (V_{β}) domains, which confer a preferential reactivity to self antigens^{2–4}, the existence of the latter (referred to as positive selection or MHC restriction) has so far only been inferred indirectly from functional studies^{5–8}. We show here that intrathymic deletion of V_{β}^{+} T cells (reactive with a self-antigen encoded by the Mls^a locus) is controlled by polymorphic MHC class II determinants. Furthermore, in mice lacking expression of Mls^a , the same class II MHC loci control the frequency of occurrence of V_{β}^{+} cells among mature CD4⁺ T lymphocytes. These data are direct evidence for positive selection by MHC determinants in the thymus in unmanipulated animals.

Recent studies have established that certain T-cell antigen receptor (TCR) V_{β} domains confer preferential reactivity *in*

Table 1 Elimination of V_{β}^{+} cells in $Mls^{a \times b}$ mice is MHC class II restricted

Strain	Mls	MHC of B10 congenic parent				% Positive T cells	
		K	A	E	D	$V_{\beta 6}$	$V_{\beta 8.2}$
DBA/1	a	–	–	–	–	6.0	12.4
(DBA/1 × B10) _{F1}	a × b	b	b	–	b	2.1	14.3
(DBA/1 × B10.D2) _{F1}	a × b	d	d	d	d	0.6	15.5
(DBA/1 × B10.S) _{F1}	a × b	s	s	–	s	4.9	11.5
(DBA/1 × B10.BR) _{F1}	a × b	k	k	k	k	0.5	17.3
(DBA/1 × B10.G) _{F1}	a × b	q	q	–	q	6.2	12.2
(DBA/1 × B10.AQR) _{F1}	a × b	q	k	k	d	0.5	18.4
(DBA/1 × B10.T.6R) _{F1}	a × b	q	q	–	d	6.4	11.4

Cortisone-resistant thymocytes from pools of 2–3 mice were stained with monoclonal antibodies (mAbs) 44-22-1 (ref. 22) (anti- $V_{\beta 6}$) (ref. 23) or F23.2 (anti- $V_{\beta 8.2}$) (ref. 3), followed by fluoresceinated goat anti-rat immunoglobulin (Tago, Burlingame, California) or goat anti-mouse IgG₁ (Southern Biotechnology, Alabama) respectively. Samples were analysed on a FACS II flow cytometer (Becton and Dickinson) gated to exclude non-viable cells²⁴. Data are presented as per cent positive cells (following subtraction of background staining with the fluorescent conjugate alone).

vitro to antigens encoded by the class II MHC (ref. 2) or minor lymphocyte stimulatory (Mls) (refs 3 and 4) loci. Furthermore, selective elimination of T cells bearing these TCR was observed in animals expressing the appropriate MHC or Mls^a gene product, which is evidence for a clonal deletion mechanism of tolerance to self-antigens. As shown previously for $V_{\beta 8.1}$ (ref. 3), intrathymic deletion of $V_{\beta 6}$ mature T cells in $Mls^{a \times b}$ mice is dependent on expression of the appropriate MHC gene products (Table 1). Thus, effective deletion occurs in congenic H-2^k or H-2^d mice, whereas H-2^b and H-2^s mice are less efficient and H-2^q mice show little or no deletion. Formal genetic mapping of this deletion phenomenon to MHC class II gene products (*I-A* and/or *I-E*) was achieved by comparing congenic B10.AQR (A^kE^k) and B10.T (6R) (A^qE^q) mice (Table 1). Elimination was selective for $V_{\beta 6}$ because T cells expressing $V_{\beta 8.2}$ (which is not normally associated with Mls^a reactivity)³ were present at similar levels, irrespective of MHC haplotype (Table 1). These data thus imply that $V_{\beta 6}$ TCR recognize Mls^a gene products in association with MHC class II determinants in a broadly polymorphic but MHC-restricted fashion.

The observed MHC restriction pattern of clonal deletion of $V_{\beta 6}$ cells *in vivo* in Mls^a mice is in accord with previous *in vitro* reports of Mls^a -reactive T-cell clones and hybrids. In the latter studies^{9–12}, Mls^a reactivity could generally be demonstrated using stimulating cells of H-2^d, H-2^k or (to a lesser extent) H-2^b genetic backgrounds, but not for H-2^q. An example of such analysis using a $V_{\beta 6}$ Mls^a -specific T-cell hybrid (RG17.16) is illustrated in Fig. 1. It can be seen that RG17.16 recognizes Mls^a (as evidenced by interleukin-2 production) on splenic stimulator cells expressing H-2^k but not H-2^q. Furthermore, this recognition could again be formally mapped to MHC class II gene products (Fig. 1).

It is of interest that optimal Mls^a reactivity of $V_{\beta 6}$ T cells (both *in vivo* and *in vitro*) correlates with cell-surface expression of the *I-E* molecular complex by appropriate stimulating cells. In mouse strains lacking *I-E* expression (E^q) (ref. 13), such stimulating (or deleting) capacity is either diminished (H-2^b, H-2^s) or undetectable (H-2^q). Interpretation of this correlation is complicated in view of the fact that the Mls^a gene product has not been defined; however, the data raise the possibility that $V_{\beta 6}$ TCR could interact directly with Mls^a and/or *I-E* molecules. If TCR interaction with *I-E* were to occur, it might be expected to influence the mature T-cell repertoire, as even low-affinity TCR interactions would be expected to lead (at least in some instances) to positive selection events¹.

Table 2 Positive selection of CD4⁺ V_{β6} T cells controlled by MHC class II gene products

Congenic strain	MHC locus				CD4 ⁺		CD8 ⁺	
	K	A	E	D	V _{β6}	V _{β8.2}	V _{β6}	V _{β8.2}
B10	b	b	-	b	7.2 ± 0.3 (3)	11.2, 10.5	9.2, 9.7	13.7
B10.BR	k	k	k	k	8.8 ± 0.3 (6)	12.2 ± 0.3 (4)	13.6 ± 1.1 (4)	14.1, 10.5
B10.D2	d	d	d	d	9.3 ± 0.2 (5)	11.8 ± 0.6 (4)	12.4 ± 1.1 (3)	19.9, 14.9
B10.G	q	q	-	q	3.8 ± 0.2 (5)	13.7 ± 0.9 (4)	8.6 ± 0.5 (4)	12.2, 9.7
B10.S	s	s	-	s	6.4 ± 0.2 (5)	12.1 ± 0.5 (4)	11.0 ± 0.3 (3)	10.8, 9.6
B10.M	f	f	-	f	6.8	16.4	8.4	8.5
B10.GD	d	d	-	b	7.1, 6.6	11.5, 11.6	10.4, 9.8	11.8, 13.0
B10.A(4R)	k	k	-	b	6.2, 6.5	14.6, 14.3	12.3, 13.2	9.1, 8.1
B10.S(9R)	s	s	k	d	10.3, 10.4	13.7, 14.0	9.9, 10.5	9.7, 8.6
B10.AQR	q	k	k	d	9.1 ± 0.2 (4)	13.3 ± 0.2 (3)	12.6 ± 1.1 (3)	13.2 ± 1.2 (3)
B10.T(6R)	q	q	-	d	4.9 ± 0.1 (4)	13.5, 12.0	8.6 ± 0.5 (3)	8.6, 8.1
(B10.G × B10.BR)F ₁	q/k	q/k	-/k	q/k	9.5 ± 0.2 (3)	15.0 ± 0.3 (3)	13.1 ± 0.5 (3)	13.4 ± 0.3 (3)

Most data were obtained by two-colour immunofluorescence of CRT stained with anti-CD8 mAb 35-17.2 and either 44-22-1 (anti-V_{β6}) or F23.2 (anti-V_{β8.2}) mAbs (see Fig. 2 for example). In some experiments CD8-depleted CRT (obtained by treatment with IgM anti-CD8 mAb 3.168.1 (ref. 28) plus complement) were used as a source of CD4⁺ cells. Data are presented as per cent positive cells (relative to total CD4⁺ or CD8⁺ subset). Mean ± s.e.m. is indicated where the number of experiments (in parentheses) was ≥ 3; otherwise data from independent experiments are separated by a comma.

To test the postulate that V_{β6} TCR may be positively selected by MHC class II (and particularly I-E) gene products (in the absence of *Mls^a*), we compared expression of V_{β6} in MHC congenic B10 (*Mls^b*) mice. For these experiments, cortisone-resistant thymocytes (CRT) were used as a source of mature T lymphocytes to avoid any skewing of the repertoire as a result of exposure to environmental antigens. Furthermore, CD4⁺ and CD8⁺ T-cell subsets were analysed independently to detect any small bias that might otherwise be masked. As shown in Fig. 2 and Table 2, reproducible differences in V_{β6} expression could be detected among CRT from B10 congenic mice. Thus V_{β6} CD4⁺ T cells were significantly (2-3-fold) more frequent in H-2^k and H-2^d haplotypes as compared with H-2^q, whereas intermediate values were obtained for H-2^b, H-2^f and H-2^s. This apparent bias towards greater V_{β6} expression in E⁺ mice was confirmed in a series of B10 recombinant inbred strains (Table 2): V_{β6} CD4⁺ cells were more frequent in B10.S(9R) mice (*A^sE^k*) than in B10.GD (*A^dE⁻*) or B10.A(4R) (*A^kE⁻*). Moreover, the frequency of V_{β6} CD4⁺ cells was controlled by MHC class II gene products (compare B10.AQR and B10.T(6R)). Evidence that this repertoire skewing is specific for V_{β6} was provided by the comparable levels of CD4⁺ T cells expressing V_{β8.2} that are independent of MHC haplotype (Table 2). The situation for the CD8⁺ subset was less clear; although we observed reduced levels of V_{β6} CD8⁺ cells in H-2^q mice,

this pattern was not confirmed in other E⁻ strains such as B10.A(4R) (Table 2).

The linkage of V_{β6} CD4⁺ T-cell frequency with expression of certain MHC class II alleles is consistent with the predictions of positive selection; nevertheless it could still be argued that negative selection (for some minor antigen distinct from *Mls^a*) accounts for the reduced levels of V_{β6} cells in H-2^q haplotypes for example. But this explanation can be excluded because V_{β6} CD4⁺ cells are as frequent in B10.G × B10.BR (H-2^q × H-2^k) F₁ animals as in the parental B10.BR strain (Table 2).

The parallel findings that both clonal elimination and overexpression of TCR using V_{β6} can be controlled by the same MHC class II genes (in the presence or absence of *Mls^a* respectively) provide compelling evidence for both a negative and a positive selection mechanism operating in the thymus. In this context, it is noteworthy that negative selection leads to deletion of both V_{β6} CD4⁺ and V_{β6} CD8⁺ T-cell subsets, whereas positive selection affects primarily V_{β6} CD4⁺ cells. The former result is compatible with recent data¹⁴⁻¹⁶ suggesting that negative selection acts on CD4⁺ thymocytes which are themselves precursors of both CD4⁺ and CD8⁺ mature T-cell subsets. With respect to positive selection, the increased frequency of V_{β6} CD4⁺ cells in E⁺ mice would be consistent with both the general association of CD4⁺ phenotype with class II MHC reactivity¹⁷ and the particular correlation of this phenotype with reactivity to *Mls^a*

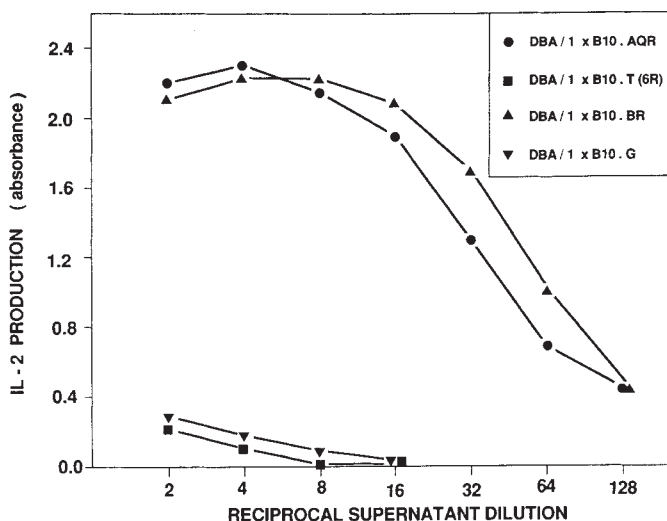


Fig. 1 MHC class II gene control of *Mls^a* stimulatory phenotype. Spleen cells from the indicated F₁ hybrid mice were used to stimulate interleukin-2 production by a V_{β6} *Mls^a*-specific T-cell hybrid (RG17.16).

Methods. RG17 was obtained by fusing a BALB/c (*Mls^b*) anti-DBA/2 (*Mls^a*) T-cell clone (AD5) (ref. 25) with the AKR thymoma BW5147. A stable subclone (RG17.16) was found to react with *Mls^a*- or *Mls^d*-bearing spleen cells. RG17.16 is CD4⁺CD8⁻ and reacts with the V_{β6}-specific mAb 44-22-1 (data not shown). For the stimulation assay, 2 × 10⁴ RG17.16 cells were cultured with 5 × 10⁵ spleen cells in 200 μl DMEM supplemented with 5% fetal bovine serum and 5 × 10⁻⁵ M 2-mercaptoethanol. After 48 h, supernatants were collected and titrated for interleukin-2 activity on CTLL-2 indicator cells. Data are presented as optical density versus supernatant concentration²⁶.

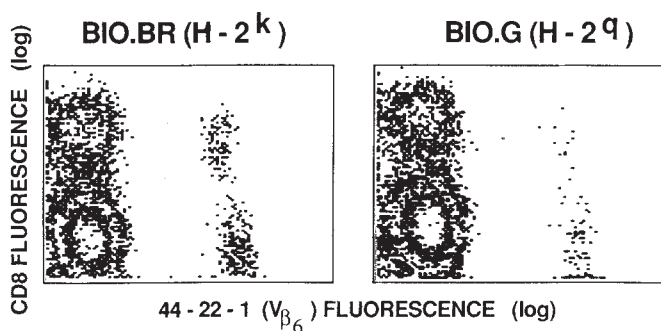


Fig. 2 Expression of $V_{\beta 6}$ by $CD8^-$ ($CD4^+$) and $CD8^+$ T-cell subsets in MHC-congenic B10 mice. CRT from B10.G (H-2^q) or B10.BR(H-2^k) mice were double-stained with mAbs directed against $V_{\beta 6}$ and CD8.

Methods. CRT from 4–6 week-old mice were obtained after a single intraperitoneal injection of hydrocortisone acetate (4 mg/per mouse) 48 h before death. All staining was at 4 °C. CRT (5×10^5 – 10^6 per sample) were incubated with undiluted hybridoma culture supernatant of mAb 44-22-1 (anti- $V_{\beta 6}$) followed by fluoresceinated goat anti-rat immunoglobulin, biotinylated mAb H35-17.2 (ref. 27) (anti-CD8) and finally phycoerythrin-conjugated avidin. Samples were analysed on a modified FACS II flow cytometer gated to exclude non-viable cells²⁴. Fifty thousand cells were accumulated using logarithmic fluorescence amplification for each cytogram. Control experiments established that >98% of CRT were either $CD4^+$ or $CD8^+$ (data not shown).

(ref. 18). Irrespective of the precise interpretation, our data suggest that TCR using certain V_{β} domains (such as $V_{\beta 6}$) have an intrinsic affinity for some polymorphic region of the MHC class II molecule, which is consistent with the postulate of Jerne¹⁹ and with more recent TCR/MHC-modelling studies²⁰. Whether positive selection occurs as a result of reactivity of such TCR with 'naked' MHC class II molecules or with class II molecules bound to a (putatively) random assortment of self-peptides²¹ cannot be answered in the absence of more precise data concerning the nature of thymic non-lymphoid cells responsible for the selective event.

We thank Dr A. Glasebrook for providing the RG17.16 cell line, Pierre Zaech for flow cytometry and Anne-France Brunet and Josiane Duc for preparing the manuscript. This work was supported in part by a grant (to H.H.) from the Swiss National Science Foundation.

Thymic selection process induced by hybrid antibodies

F. Zepp* & U. D. Staerz†

Basel Institute for Immunology, Grenzacherstrasse 487, CH-4005 Basel, Switzerland

Thymus-derived (T) lymphocytes using the $\alpha\beta$ T-cell antigen receptor (TCR) recognize fragmented antigen in conjunction with surface molecules encoded by genes of the major histocompatibility complex (MHC)^{1–6}. Peripheral T lymphocytes preferentially see antigen presented by self rather than by foreign MHC molecules^{7–9}, and autoreactive T lymphocytes are deleted¹⁰. Thus, the peripheral T-lymphocyte repertoire is skewed towards recognition of antigen in the context of self-MHC and towards tolerance to self-antigens. During T-lymphocyte development in the thymus, this repertoire is formed by the interaction of TCR with MHC molecules resulting in positive and negative selection phenomena^{7–10}. Hybrid antibodies (HABs) that carry binding sites to the TCR and to a surface marker on another cell can engage all T lymphocytes regardless of their specificity¹¹. It should be possible to mimic selection processes in normal animals with HAB that specifically link members of a TCR family to MHC molecules on the thymic stroma. We have probed T-lymphocyte development with HABs linking $V_{\beta 8}$ -positive TCR to either class I or class II MHC products in thymic organ culture. Thymocytes exposed to either HAB in an early stage of maturation respond with a significant increase in the frequency of $V_{\beta 8}$ -carrying cells. At a later stage of development $V_{\beta 8}$ -positive thymocytes are depleted. These results illustrate the succession of positive and negative selection in the developing thymus of normal mice.

Thymocytes in fetal organ culture develop in an apparently normal manner^{12,13}; double negative $CD4^-CD8^-$ (DN) thymocytes appear early followed by double positive $CD4^+CD8^+$ (DP) cells and the single positive, $CD4^+$ or $CD8^+$, thymocytes at a later stage¹⁴. As thymuses are populated by waves of precursors starting at about day 12 of fetal life, organs obtained at different time points should allow access to thymocytes at different maturation stages¹⁵. However, due to lack of functional assays for non-mature T-lymphocyte precursors, it has been impossible to observe the effects of thymic selection processes on developing thymocytes. This problem could be overcome in constellations where usage of a certain TCR chain is linked to antigenic specificity. In these cases, effects of negative selection in the thymus could be directly followed^{16–18}. By creating TCR transgenic mice it could be demonstrated that clonal deletion acts upon non-mature DP thymocytes¹⁹. The mechanisms that operate in positive selection and their connection with negative

* Permanent address: Department of Paediatrics, Mainz University, Mainz, FRG.

† To whom correspondence should be addressed.

Received 14 July; accepted 19 October 1988.

- Moller, G. *Immunol. Rev.* **101**, 1–215 (1988).
- Kappler, J. W., Roehm, N. & Marrack, P. *Cell* **49**, 273–280 (1987).
- Kappler, J. W., Staerz, U., White, J. & Marrack, P. *Nature* **332**, 35–40 (1988).
- MacDonald, H. R. *et al.* *Nature* **332**, 40–45 (1988).
- Bevan, M. J. *Nature* **269**, 417–418 (1977).
- Zinkernagel, R. M. *et al.* *J. exp. med.* **147**, 882–896 (1978).
- Marrack, P., Kushnir, E., Born, W., McDuffie, M. & Kappler, J. *J. Immun.* **140**, 2508–2514 (1988).
- Singer, A. *J. Immun.* **140**, 2481–2483 (1988).
- Glasebrook, A. L. & Fitch, F. W. *J. exp. med.* **151**, 876–895 (1980).
- Webb, S. R., Molnar-Kimber, K., Bruce, J., Sprent, J. & Wilson, D. B. *J. exp. Med.* **154**, 1970–1974 (1981).
- Jones, B. & Janeway, C. A. *Immunogenetics* **16**, 243–255 (1982).
- Lynch, D. H. *et al.* *J. Immun.* **134**, 2071–2078 (1985).
- Mathis, D. J., Benoist, C., Williams, V. E., Kanter, M. & McDevitt, H. O. *Proc. natn. Acad. Sci. U.S.A.* **80**, 273–277 (1983).
- Kisielow, P., Bluthmann, H., Staerz, U. D., Steinmetz, M. & von Boehmer, H. *Nature* **333**, 742–746 (1988).
- Fowlkes, B. J., Schwartz, R. H. & Pardoll, D. M. *Nature* **334**, 620–623 (1988).
- MacDonald, H. R., Hengartner, H. & Pedrazzini, T. *Nature* **335**, 174–176 (1988).
- Swain, S. L. *Immunol. Rev.* **74**, 129–142 (1983).
- Janeway, C. A. & Katz, M. E. *J. Immun.* **134**, 2057–2063 (1985).
- Jerne, N. K. *Eur. J. Immun.* **1**, 1–9 (1971).
- Davis, M. M. & Bjorkman, P. J. *Nature* **334**, 395–402 (1988).
- Kourilsky, P. & Claverie, J.-M. *Annls Inst. Pasteur, Paris* **137D**, 3–21 (1986).
- Acha-Orbea, H., Zinkernagel, R. M. & Hengartner, H. *Eur. J. Immun.* **15**, 31–36 (1985).
- Payne, J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- MacDonald, H. R. & Zaech, P. *Cytometry* **3**, 55–58 (1982).
- Kelso, A., Glasebrook, A. L., Kanagawa, O. & Brunner, K. T. *J. Immun.* **129**, 550–556 (1982).
- Landegren, U. *J. Immunol. Meth.* **67**, 379–388 (1984).
- Golstein, P. *et al.* *Immunol. Rev.* **68**, 5–42 (1982).
- Sarmiento, M. *et al.* *Immunol. Rev.* **68**, 135–169 (1982).

Table 1 Frequency of $V_{\beta 8}$ expression in early thymus cultures

Treatment	TCR		% of $CD3^+$ thymocytes expressing $V_{\beta 8}$
	$V_{\beta 8}$	CD3	
—	% of cultured thymocytes		
H5	10.8	32.2	33.5
H5	36.5	60.0	60.8
H4	34.1	56.4	60.5

Thymocytes obtained from the same cultures as described in Fig. 1a, were stained in indirect immunofluorescence for $V_{\beta 8}$ TCR expression using F23.1 and/or CD3 expression using 500 A2.

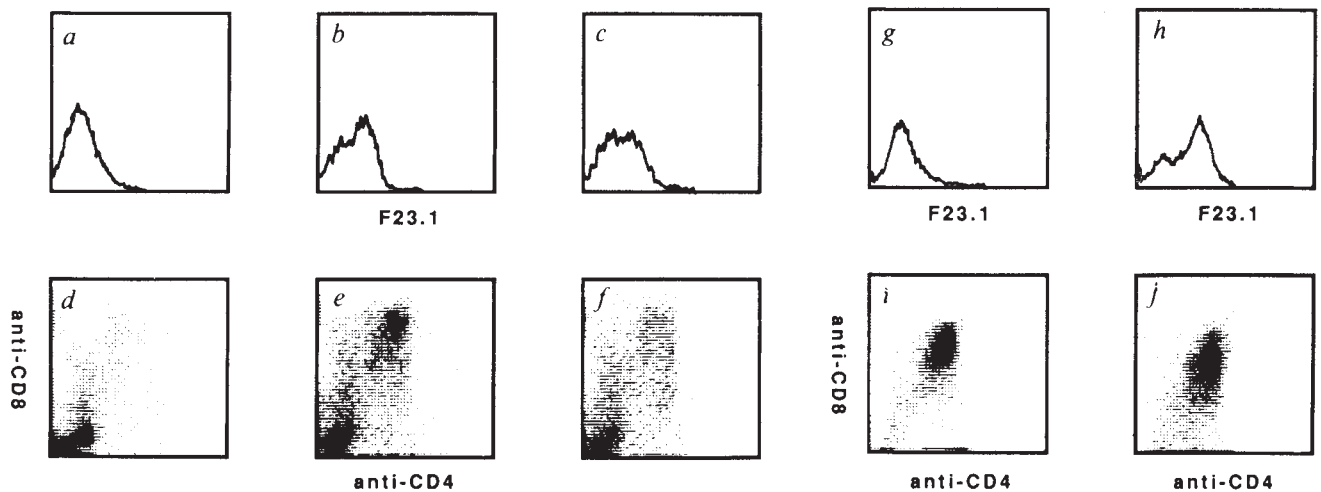


Fig. 1 Influence of HAb on the expression of $V_{\beta}8$ -positive TCR on thymocytes of fetal day 13 (*a-f*) or day 14 (*g-j*) thymic cultures. Thymuses collected from DBA/2 fetuses were cultured with no antibodies (*a, d*) (*g, i*), with H5 (*b, e, h, j*), or with H4 (*c, f*) and collected five days later. Thymocytes were stained for $V_{\beta}8$ TCR with F23.1 in indirect immunofluorescence (*a, b, c, g, h*), expression of CD4 and CD8 in two-colour immunofluorescence (*d, e, f, i, j*).

Methods. A HAT-sensitive sub-line of F23.1^{24,25} was transformed to neomycin-resistance with the retrovirus N2, kindly provided by Dr G. Keller²⁶, and fused with the B hybridomas 15.5.5 and 14.4.4 that recognize H-2D^{d/k} and H-2E^{d/k}, respectively. Supernatants from HAb H5 (F23.1-15.5.5) and H4 (F23.1-14.4.4) were enriched for the antibody by affinity chromatography on Protein A Sepharose (Pharmacia) and the HAb were purified by repetitive runs over a Bio Gel HPHT column (Bio-Rad)²⁷. Embryonic thymuses obtained from DBA/2 fetuses of defined age were introduced into organ culture in Iscove's modified Dulbecco's medium supplemented with FCS, non-essential amino acids and antibiotics. Cultures were established in the presence of $10 \mu\text{g ml}^{-1}$ of purified HAb, H5 or H4. On day four of culture, the organs were washed with three changes of medium and incubated for an additional 24 hours to allow capping of bound antibodies, as we had determined in preliminary experiments. On day five, single-cell suspensions were prepared and $3-4 \times 10^5$ thymocytes per lobe could be recovered. For staining in PBS/5% FCS the following monoclonal antibodies were used: F23.1, biotinylated 53.6.75 (anti-CD8), and fluorescein isothiocyanate-conjugated GK1.5 (anti-CD4) prepared by standard procedure. For indirect immunofluorescence, cells were incubated with F23.1 for 20 min at 37°C followed by fluorescein isothiocyanate-labelled rabbit anti-mouse immunoglobulin (Silenus) for 20 min at 4°C after several washes with medium. The x axis shows relative fluorescence and the y axis relative cell numbers. For two-colour fluorescence, cells were incubated with biotinylated 53.6.75 at 37°C , washed twice and stained with fluorescein isothiocyanate-conjugated GK1.5 in a mixture together with phycoerythrin streptavidin (Becton Dickinson) at 4°C . The extent of staining was analysed on a FACScan flow-cytometer (Becton Dickinson) with a single argon laser and logarithmic intensity using FACScan research software program (FRSP). In each sample, 7,500 viable cells were analysed. The results are presented as 'density' plots, generated by analysis of processed data reduced to a 64×64 matrix with sixteen levels. The x axis shows green fluorescence, the y axis red fluorescence. A sector of $4 \times 4 \log_{10}$ units of relative fluorescence is depicted. Percentages of stained and non-stained cells were calculated using FRSP.

selection still remained obscure. To gain insight into the sequence of thymic education, we decided to determine the fate of maturing thymocytes expressing $V_{\beta}8$ -positive TCR, a marker that is only loosely connected with antigenic specificity. Triggering of T-lymphocytic precursors would be achieved by HAb that link $V_{\beta}8$ -positive TCR to molecules on the thymic stroma, in close approximation to the physiological situation, to MHC molecules. As a stable source of native HAb, hybrid hybridomas were established by fusion of a hypoxanthine, aminopterin and thymidine (HAT)-sensitive neomycin-resistant subline of F23.1 (anti-TCR $V_{\beta}8$) with either of two HAT-resistant hybridomas secreting monoclonal antibodies reactive with class I or class II products of the d and k, but not b, haplotype of murine MHC. Thus, H5 secretes HAb with a binding site for any of the three members of the TCR $V_{\beta}8$ family and a binding site for H-2D^{d/k} molecules. H4 was constructed to link $V_{\beta}8$ -positive TCR to H-2E^{d/k} molecules.

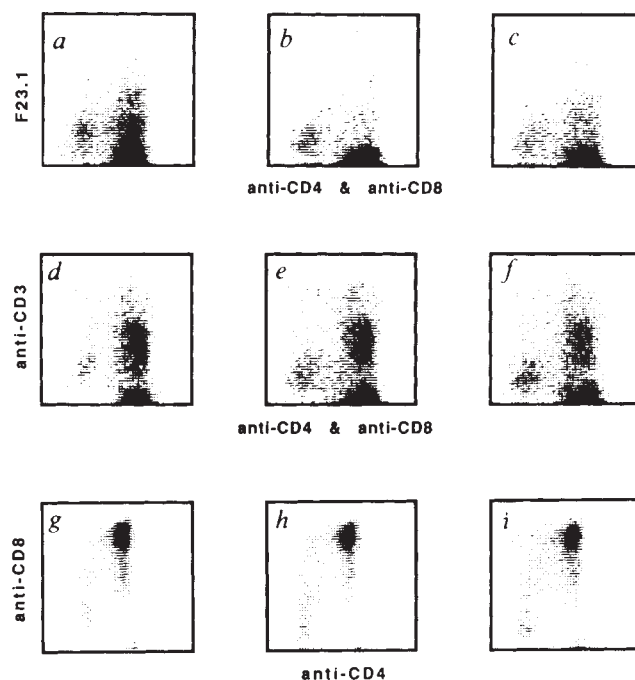
In the first set of experiments, thymuses were collected from day 13 DBA/2 (H-2^d) fetuses as a source of thymocytes early in development. After five days of culture, most thymocytes in the control group had not developed beyond the immature DN stage (Fig. 1*d*). The introduction of purified HAb, H5 or H4, drastically altered this distribution (Fig. 1*b, c, e, f*). The overall frequency of $V_{\beta}8$ -positive thymocytes increased more than threefold (Table 1). Simultaneously, cells of a more mature phenotype appeared. In the case of H5, and of H4, the frequency of DP cells increased from 12.1% to 32.3%, roughly equal to the accumulation of $V_{\beta}8$ -positive thymocytes. We could further

determine that all newly appearing DP thymocytes express $V_{\beta}8$ -positive TCR (data not shown). Therefore, these changes in the development were directed by the specificity of the HAb. In cultures established from day 14 DBA/2 fetal thymuses, the majority of cells had already advanced to the DP phenotype (Fig. 1*g, h, i, j*). In this situation, H5 produced an even more pronounced increase in the frequency of $V_{\beta}8$ -positive DP thymocytes that seemed blast-like in size and carried the J11d marker, indicating that they were at an immature stage of development (data not shown). No change in $V_{\beta}8$ expression was seen in DN or single-positive thymocytes in all early cultures. Controls with age-matched fetuses were set up: DBA/2 thymuses were exposed to unlinked parental antibodies, and C57BL/6 (H-2b) thymuses were cultured with HAb. Under these conditions the antibodies did not show any significant influence (data not shown). H5 can indeed link TCR to class I molecules on any cell, but due to the distribution of the MHC molecules in the murine system, H4 is restricted to attach the TCR to class II molecules in thymic stroma. Thus, we conclude that interaction of TCR with the thymic stroma is the most likely triggering event, although the possibility that class II molecules had been picked up by thymocytes²⁰ cannot be excluded. To examine whether interactions of TCR and MHC molecules are prerequisites for thymocyte maturation was beyond this study.

Attempts to follow the development of positive selected cells for more than five days in this culture system resulted in high levels of cell death making interpretation impossible¹². Therefore, we resorted to using more mature (day 16) fetuses as a

Fig. 2 Comparison of cell-surface expression of $V_{\beta}8$ -positive TCR, CD4 and CD8 on thymocytes of fetal day 16 thymic organ culture from CBA/2 animals supplemented with HAb. Thymocytes were collected from organs cultures without antibodies (a, d, g) or in the presence of H5 (b, e, h), or H4 (c, f, i). The cells were analysed for expression of $V_{\beta}8$ -positive TCR (a, b, c) or CD3 (d, e, f) in comparison to staining for CD4 together with CD8, or expression of CD8 in comparison to CD4 (g, h, i) by two-colour flow cytometry.

Methods. The staining and analysis was performed as described in detail in Fig. 1, using the following additional monoclonal antibodies: biotinylated F23.1, biotinylated 500 A2²⁸, and fluorescein isothiocyanate-conjugated 53.6.75.



source of thymuses at an advanced stage of development. CBA/CA (H-2^k) animals were chosen to circumvent tolerance phenomena acting upon the $V_{\beta}8.1$ member of the TCR $V_{\beta}8$ family in Mls^a mice¹⁷. In contrast to the early cultures, we observed that HAb reduced the cellular yield by approximately 20%, whereas the size distribution remained unaltered. Statistical evaluation of the two-colour staining in Fig. 2 showed that both HAb caused a significant shift (approximately 6%) from the DP to the DN phenotype. A depletion (between 16.9% and 9.2%) of thymocytes expressing $V_{\beta}8$ -positive TCR with a corresponding reduction of CD3 positive cells was also seen (Table 2). As anti-TCR antibodies bypass the requirements for accessory molecules, like CD8 and CD4, it could not be expected that H5 and H4 specifically deleted precursors of mature single-positive thymocytes depending on the MHC targeted. Accordingly, both HAb caused an equal reduction (~1%) in both single-positive populations. Therefore, the major reduction of $V_{\beta}8$ positive cells takes place within the non-mature, DP subset, rather than within mature thymocytes. The extent of specific deletion was unexpected because also in these 'late' cultures prothymocytes should give rise to TCR-bearing cells that could be amplified. And, indeed, evidence for an expansion of $V_{\beta}8$ -positive cells was observed. The DN population showed an increase of $V_{\beta}8$ expression (approximately 70%) when calculated as a fraction of TCR-carrying cells (Table 2). This subpopulation might represent the late appearing, non-functional J11d⁻ DN cells that express TCR^{21,22}. They could be remnants of cells that had gone through both selection steps. In control experiments using C57BL/6 mice, only a minor decrease in the high-TCR-expressing cells was noted. A more pronounced effect was

induced by a mixture of the parental antibodies in the CBA/CA thymuses as had been described elsewhere²³.

These results show thymic education as a succession of positive and negative selection events during thymic development. DP thymocytes were identified as the developmental stage when the T-lymphocyte repertoire of antigenic specificities is shaped. HAb targeting thymocytes to different cell surface structures will help to determine the importance of the MHC molecules for thymic education and answer questions regarding the function of different cellular components of the thymic stroma.

We would like to thank Drs H. von Boehmer, D. Gray, K. Karjalainen, P. Kisielow, P. Matzinger and B. Stockinger for critical reading and Ms J. Millar for typing of the manuscript. The Basel Institute for Immunology was founded and is supported by F. Hoffman-La Roche & Co. Ltd, Basel, Switzerland.

Received 27 June; accepted 25 October 1988.

Table 2 Surface phenotype and TCR expression in late thymus cultures

Treatment	Markers		Cell-surface phenotype			
	$V_{\beta}8$ (CD3)	—(—)	+(+)	—(—)	+(+)	+(+)
	CD4 and CD8	—(—)	—(—)	+(+)	+(+)	+(+)
—		6.7 (1.8)	4.0 (9.0)	60.8 (33.8)	28.5 (55.4)	
H5		9.4 (6.8)	6.8 (9.9)	72.2 (41.5)	11.6 (41.8)	
H4		9.8 (5.6)	6.7 (9.2)	64.2 (37.8)	19.3 (47.8)	

Two-colour immunofluorescence data from Fig. 2 were calculated using FRSP software (Fig. 1 legend). Values in brackets show staining for CD3 instead of F23.1.

- Kindred, B. & Shreffler, D. C. *J. Immun.* **109**, 940-943 (1972).
- Katz, D. H., Hamaoka, T. & Benacerraf, B. *J. exp. Med.* **137**, 1405-1418 (1973).
- Zinkernagel, R. M. & Doherty, P. C. *J. exp. Med.* **141**, 1427-1436 (1975).
- Babbitt, B., Allen, P., Matsueda, G., Haber, E. & Unanue, E. *Nature* **317**, 359-361 (1985).
- Schwartz, R. A. *Rev. Immun.* **3**, 237-261 (1985).
- Townsend, A. R. M., Gotch, F. M. & Davey, J. *Cell* **42**, 457-467 (1985).
- Bevan, M. J. *Nature* **269**, 417-419 (1977).
- von Boehmer, H., Haas, W. & Jerne, N. K. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2439-2442 (1977).
- Zinkernagel, R. & Doherty, P. C. *Adv. Immun.* **27**, 51-77 (1987).
- Nossal, G. J. V. *Rev. Immun.* **1**, 33-62 (1983).
- Staerz, U. D., Kanagawa, O. & Bevan, M. J. *Nature* **314**, 628-631 (1985).
- Robinson, J. H. & Owen, J. J. T. *Clin. exp. Immun.* **23**, 347-354 (1976).
- Robinson, J. H. & Owen, J. J. T. *Clin. exp. Immun.* **27**, 322-327 (1977).
- Kisielow, P., Leiserson, W. & von Boehmer, H. *J. Immun.* **133**, 1117-1123 (1984).
- Snodgrass, R. & Keller, G. *EMBO J.* **6**, 3955-3960 (1987).
- Kappler, J. W., Roehm, N. & Marrack, P. C. *Cell* **49**, 273-280 (1987).
- Kappler, J. W., Staerz, U. D., White, J. L. & Marrack, P. C. *Nature* **332**, 35-40 (1988).
- MacDonald, H. R. *et al. Nature* **332**, 40-45 (1988).
- Kisielow, P., Blüthmann, H., Staerz, U. D., Steinmetz, M. & von Boehmer, H. *Nature* **333**, 742-746 (1988).
- Sharrow, S. O., Mathieson, B. J. & Singer, A. *J. Immun.* **126**, 1327-1335 (1981).
- Singer, P. A., McEvilly, R. J., Noonan, D. J., Dixon, F. J. & Theophilopoulos, A. N. *Proc. natn. Acad. Sci. U.S.A.* **83**, 7018-7022 (1986).
- Crispe, I. N. *et al. Nature* **329**, 336-339 (1987).
- Born, W. *et al. J. Immun.* **138**, 999-1008 (1987).
- Staerz, U. D., Rammensee, H.-G., Benedetto, J. & Bevan, M. J. *J. Immun.* **134**, 3994-4000 (1985).
- Behlke, M. A. *et al. J. exp. Med.* **165**, 257-262 (1987).
- Keller, G., Paige, C., Gilboa, E. & Wagner, E. F. *Nature* **318**, 149-154 (1985).
- Staerz, U. D. & Bevan, M. J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1453-1457 (1986).
- Havran, W. L. *et al. Nature* **330**, 170-173 (1987).

Antigen presenting function of class II MHC expressing pancreatic beta cells

James Markmann*, David Lo†‡, Ali Naji*,
Richard D. Palmiter§, Ralph L. Brinster†
& Ellen Heber-Katz||

* Department of Surgery, School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

† Laboratory of Reproductive Physiology, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

§ Howard Hughes Medical Institute, Department of Biochemistry, University of Washington, Seattle, Washington 98195, USA

|| Wistar Institute of Anatomy and Biology, Philadelphia, Pennsylvania 19104, USA

Class II major histocompatibility complex (MHC) gene expression in the mouse is generally limited to thymic epithelium and bone marrow-derived cells such as B lymphocytes and cells of the macrophage/dendritic cell lineage (Mφ/DC). Class II-bearing B lymphocytes and Mφ/DC possess antigen presenting cell (APC) function; that is, they can stimulate T lymphocytes reactive to either antigen plus MHC or foreign MHC alone. To assess whether non-bone-marrow-derived cells can acquire APC function and elicit graft rejection through expression of class II, we studied transgenic pancreatic islet beta cells that express a foreign class II (I-E) molecule. *In vivo*, grafts of I-E⁺ transgenic islets into I-E⁻ naive hosts are not rejected unless the host is primed by an injection of I-E⁺ spleen cells. *In vitro*, the I-E⁺ beta cells are unable to stimulate T lymphocytes reactive to I-E plus a peptide antigen. Paradoxically, they induce antigen specific unresponsiveness in the T cells. We propose that expression of class II on non-lymphoid cells may serve as an extrathymic mechanism for maintaining self tolerance.

Receptors on class II-restricted T lymphocytes recognize a compound determinant comprised of a processed peptide antigen and a class II MHC heterodimer¹⁻³. Presentation of this complex by APC is critical to the activation of T cells, but it is not completely understood what other features are important for effective stimulation of resting T cells. Clearly, variability exists in the efficiency of antigen presentation by class II-expressing bone marrow derived cells; Mφ/DC appear to be as much as 10,000-fold more potent than resting B cells^{4,5}. Despite these findings, it has been assumed that nearly any cell can acquire APC function through class II expression induced by lymphokines such as γ -IFN and tumour necrosis factor^{6,7}. Thus, it has been proposed that parenchymal cells of various endocrine organs, induced to express class II, present self antigens to potentially autoreactive T cells and induce autoimmune disease⁸. Consistent with this hypothesis are the observations that pancreatic islet beta cells from patients with insulin-dependent diabetes mellitus and thyrocytes undergoing attack in autoimmune thyroiditis aberrantly express class II antigens⁹.

In a previous report¹⁰, we described studies addressing this hypothesis using transgenic mice (Ins-I-E) which expressed class II I-E^b (E_A^d/E_B^b) molecules on beta cells of the pancreatic islet, but not on bone marrow derived cells. In these studies, transgenic islets were found to have decreased ability to produce insulin, even though I-E⁺ cells (presumably beta cells) persisted for months (ref. 10, and D.L. *et al.*, unpublished results). This result suggested that high levels of class II MHC expression in some way impaired insulin production by beta cells. Similar studies have also been reported by others^{11,12}. Autoimmune destruction of the pancreatic islets was not observed in any of the studies.

Table 1 Ability of transgenic fetal pancreas to stimulate graft rejection *in vivo*

Exp.	Pancreas donor	Recipient priming with I-E ⁺ spleen	Grafts rejected/ total grafted			
			d12	d24	d47	d130
1	Control I-E ⁻	—		0/4		0/3
		+(d0)	0/6	0/3		
	Ins-I-E	—		0/5		0/7
		+(d0)	4/4	3/3		
2	Control I-E ⁻	+(d40)			0/6	
	Ins-I-E	+(d40)			4/4	
3	Control I-E ⁻	—		0/5		
	Wild-type I-E ⁺	—		4/4*		
		+(d0)	6/6*			

Fetal pancreas grafts (19-day gestation) from B6 females mated to either Ins-I-E (line 187-7) or wild-type I-E (line 107-1; ref. 10) transgenic males were transplanted under the kidney capsule of male (B6×SJL)F₁ recipients. Tissue from fetal pancreas donors was analysed for the presence of the transgene²¹. Non-transgenic littermates served as negative control grafts. Where indicated, recipient mice were immunized with 20×10⁶ wild-type I-E⁺ spleen cells subcutaneously on the day of transplantation or 40 days post grafting. Graft survival was evaluated histologically at times indicated, scored blind by a pathologist. Rejection was defined as lymphocytic infiltration of grafts (hematoxylin-eosin-stained sections) and/or depletion of insulin granule-containing cells (aldehyde fuchsin, which specifically stains mature insulin granules). It should be noted that although insulin staining was weaker in transgenic islets¹⁰, weak insulin staining persisted in grafts even 130 days post grafting. This suggested that beta cells survived throughout the study period, even though insulin production by these cells is diminished. * Histological rejection of wild-type I-E⁺ grafts differed from rejection of Ins-I-E grafts. Although significant infiltration was observed, islets frequently remained morphologically intact (Fig. 1g), and insulin granules were still present at the times observed. This difference may reflect the fact that wild-type I-E⁺ mice do not express I-E on islet cells.

Furthermore, T cells from Ins-I-E mice were tolerant to the transgene I-E even though I-E expression was absent in thymus or spleen¹⁰. To analyse this phenomenon in more detail we isolated the I-E expressing transgenic tissues for *in vivo* and *in vitro* studies to determine whether the APC function of the beta cells was relevant to the tolerant phenotype.

For antigen presentation *in vivo*, we examined the response of naive I-E⁻ mice to fetal pancreas transplanted from Ins-I-E mouse donors. In these experiments, although whole fetal pancreas is grafted, the exocrine tissue atrophies, leaving endocrine islet tissue intact (Fig. 1 and refs 13, 14). Fetal pancreas grafts from Ins-I-E transgenic mice (line 187-7) or their non-transgenic littermates were accepted indefinitely (Table 1, experiment 1, Fig. 1a–d). Fetal pancreas grafts from wild-type I-E⁺ mice were grafted into another set of mice as positive controls. In this case, the grafts contained I-E⁺ Mφ/DC, and the grafts were rejected (Table 1, experiment 3, but see footnote to Table 1 and Fig. 1g). To determine whether the absence of an immune response to Ins-I-E pancreas grafts was due to an inability of I-E to be recognized by immune cells, similar Ins-I-E grafts were transplanted to recipients that had been primed to I-E at the time of transplantation (Table 1, experiment 1) or 40 days post transplantation (Table 1, experiment 2) by injection with spleen cells from wild-type I-E⁺ mice. Under these conditions Ins-I-E transgenic grafts were rejected, whereas control I-E⁻ grafts were accepted. Histologically, transgenic grafts were infiltrated by lymphocytes and complete depletion of islet cells was observed at two weeks post-priming (Fig. 1e,f). These data indicate that the I-E expressed by beta cells of transgenic grafts is recognized by immune effector cells if the response is initiated by conventional I-E⁺ APC; class II⁺ beta cells are themselves apparently incapable of initiating an immune response by naive non-tolerant T lymphocytes *in vivo*.

‡ To whom correspondence should be addressed.

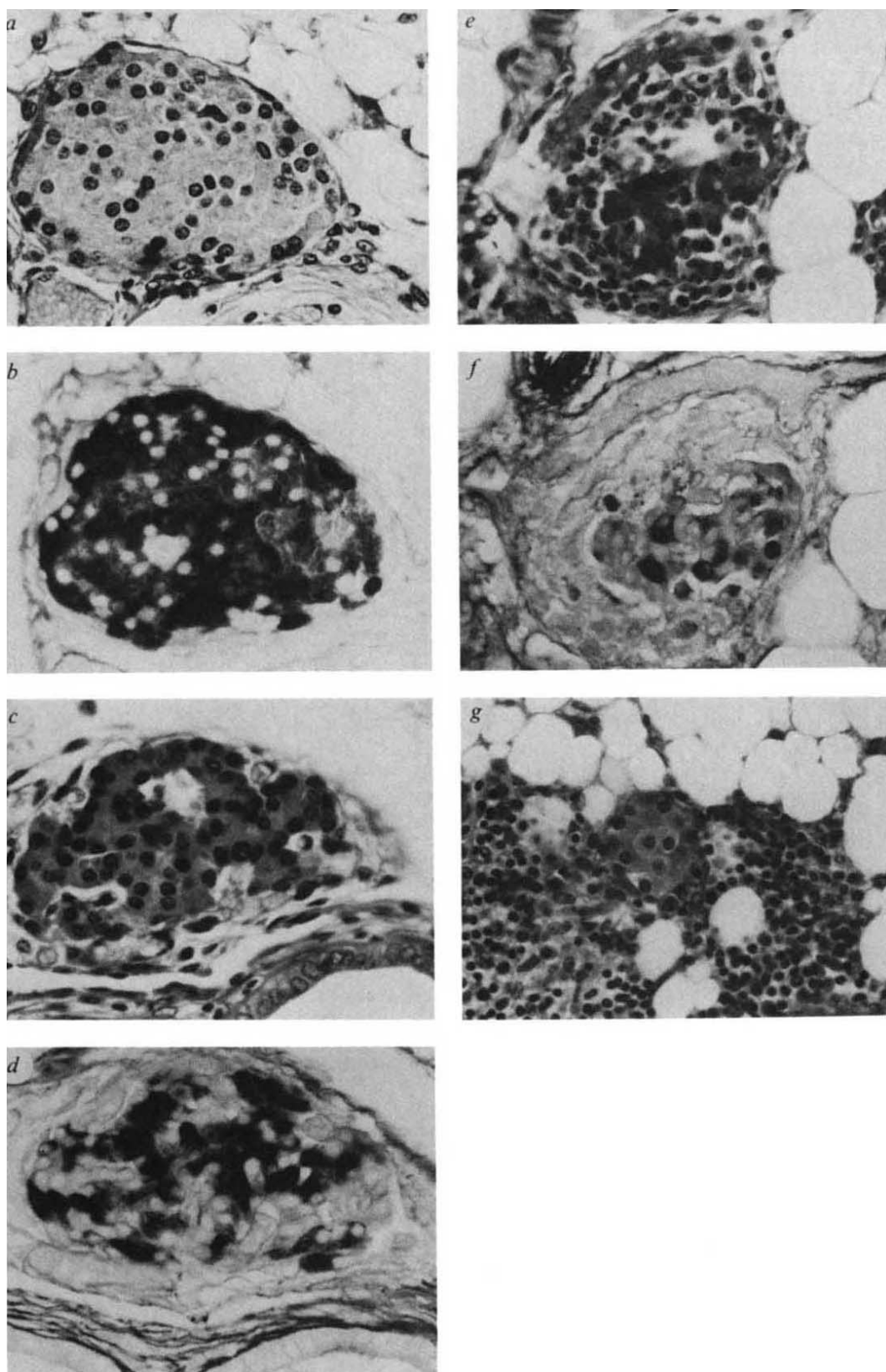


Fig. 1 Fetal pancreas grafts from non-transgenic littermate (*a, b*), and Ins-I-E donor (*c, d*) 24 days after grafting. Although whole fetal pancreas is grafted, the exocrine tissue atrophies^{13,14}, leaving endocrine islet tissue intact in both cases. In *e* and *f*, fetal pancreas graft from Ins-I-E donor, 47 days after grafting, and 7 days after the recipient was primed with I-E⁺ spleen (Table 1, experiment 2). Note the extensive lymphocytic infiltrates and islet cell destruction. In *g*, wild-type I-E⁺ fetal pancreas graft in unprimed recipient, 24 days after grafting. Infiltrates are present, but intact islet tissue is also seen. (*a, c, e* and *g*), stained with hematoxylin and eosin; (*b, d* and *f*), aldehyde fuchsin. All photographs were taken at the same magnification.

To explore further the interaction of class II-bearing beta cells with T cells, an *in vitro* system of antigen presentation was used. A B10.A(5R) (E _{α} ^k/E _{β} ^b) T-cell line specific for a peptide fragment (residues 1–23) of the herpes simplex virus glycoprotein D (gD) in the context of I-E^b was prepared as previously described¹⁵. Anti-gD specific T cells mounted a vigorous proliferative response when cultured in the presence of antigen

and irradiated B10.A(5R) spleen cell stimulators. In contrast, proliferation was not observed in response to islet cells isolated from Ins-I-E transgenic mice (both lines 187-1 and 187-7) plus antigen (data not shown). Thus, I-E⁺ beta cells from Ins-I-E mice were unable to present antigen to T lymphocytes *in vitro* despite surface expression of I-E determinants (as determined by immunofluorescence staining; Fig. 2).

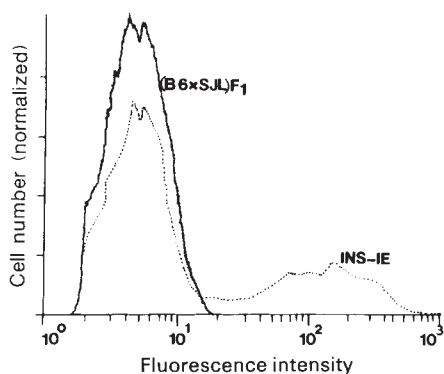


Fig. 2 Immunofluorescence analysis of Ins-I-E islet cells. Pancreatic islets were isolated from 4-week-old Ins-I-E and (B6×SJL) F_1 mice as previously described²². Islets were cultured for 24 h before use to reduce the numbers of contaminating exocrine cells. Single cell suspensions were prepared by incubation of islets in 1% trypsin for 5 min at 37 °C. Islet cells were labelled sequentially with biotinylated monoclonal antibody 14.4.4s (anti-I-E) and avidin phycoerythrin (APE) or APE alone. Islet cells from Ins-I-E and (B6×SJL) F_1 donors were 24.8 and 2.4% positive for I-E antigens respectively (mean of four experiments). I-E⁺ Ins-I-E islet cells had a mean fluorescence intensity almost double that of I-E⁺ spleen cells from B10.BR mice (data not shown).

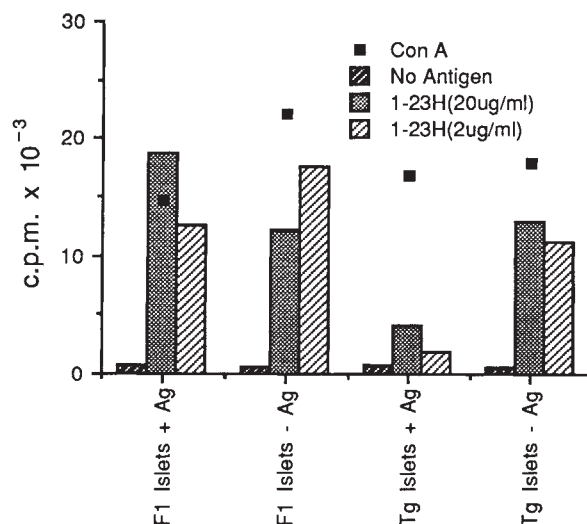


Fig. 3 Response of T-cell line to antigen on conventional APC after preculture with islet cells. B10.A(5R) mice were immunized with the 1-23(H) peptide of HSV gD and lymph node T cells removed 12 days later. A T-cell line was generated by multiply stimulating with the 1-23(H) peptide¹⁵. As indicated in the figure, T cells (10^5 per well) were added to control (B6×SJL) F_1 ('F1 Islets') or Ins-I-E (line 187-1) transgenic ('Tg Islets') islet cells (10^5 per well) in the presence or absence of 1-23(H) ($25 \mu\text{g ml}^{-1}$) in 96-well flat-bottom plates for 48-h preculture. T cells from each group were then washed six times and added (10^4 per well) to new 96-well flat-bottom plates with X-irradiated B10.A(5R) splenocytes (5×10^5 per well) plus antigen as indicated. Carryover of islet cells was not considered to be a problem, because intentional mixing of islet and spleen cells showed no evidence of suppression (data not shown). ConA was used to assess cell viability and cell number. Plates were incubated for 72 h, when [^3H]thymidine was added to measure proliferation. The figure shown is representative of four separate experiments using three different T-cell lines.

Recent studies have demonstrated that inappropriate antigen presentation to T cells, either by purified class II molecules in planar lipid membranes¹⁶ or by chemically modified APCs¹⁷, can induce a state of unresponsiveness in the exposed T cells. Considering this, it was possible that the I-E⁺ transgenic beta cells, unable to stimulate T cells either *in vivo* or *in vitro*, might instead have paralysed the responder cells. To test this, transgenic (line 187-1) and control islet cells were cultured with gD-specific B10.A(5R) T cells in the presence or absence of antigen for 48 hours. T cells recovered from these cultures were then re-challenged with antigen presented by conventional I-E⁺ B10.A(5R) spleen APCs. As shown in Fig. 3 (representative of four experiments using three different T-cell lines), a significant reduction in antigen specific T-cell reactivity was observed upon preculture in the presence of antigen and Ins-I-E transgenic islet cells, but not with transgenic islet cells alone, or with control islet cells with or without antigen. The preculture did not alter viability of the T cells, because all precultured populations responded to ConA. The response to ConA does not necessarily conflict with the apparent paralysis observed however. Signalling through the T-cell receptor is still possible in paralysed T cells^{16,17}, and proliferation of paralysed T-cell clones is still inducible by antibodies to CD3 (H. Quill, personal communication; D. Mueller, personal communication). In addition, because our studies used T-cell lines rather than clones, some T cells in the heterogeneous population (ref. 18, and E.H.-K., unpublished results) may not have been paralysed by the preculture. They would certainly respond to ConA by production of IL-2, which will induce proliferation even in paralysed cells^{16,17}.

Although we observed paralysis in three different T-cell lines *in vitro*, I-E reactive cells were not paralysed in the graft recipients *in vivo*. Indeed, T cells from naive graft recipients tolerating transgenic islet grafts for >120 days were still responsive to I-E in mixed lymphocyte cultures (data not shown). It is possible, however, that only a fraction of circulating T cells in the graft recipients were exposed to the I-E⁺ tissue, and the paralysed T cells would be obscured by the response of the remaining T cells. Furthermore, a significant proportion of I-E-reactive T cells may already have differentiated past a paralysis susceptible stage. The difference in the target antigens used (gD/I-E *in vitro*, I-E alone as alloantigen *in vivo*) is also important in that the T-cell lines *in vitro* are relatively homogeneous

in their avidity to the target antigen¹, whereas the I-E reactive T cells *in vivo* are likely to be extremely heterogeneous. Heterogeneity among the I-E-reactive T cells probably affects susceptibility to paralysis *in vivo*. It must be noted that our results differ from those reported by Lafferty and coworkers^{13,19,20}, who found reduced host reactivity to MHC antigens of allogeneic grafts tolerated for >100 days. However, aside from technical differences, these workers also used whole MHC differences rather than class II MHC alone, and that situation may introduce cooperative effects.

The present data suggest two major conclusions. First, for pancreatic islet beta cells, class II expression alone is not sufficient to endow complete APC function, and is therefore unlikely to be solely responsible for initiation of autoimmunity. The rejection of Ins-I-E grafts by primed recipients indicates, however, that beta cell class II does confer susceptibility to a conventionally initiated class II-specific immune attack *in vivo*. Second, these data provide evidence that antigen presented by class II-bearing islet cells (and perhaps most non-lymphoid cells) can result in subsequent unresponsiveness of antigen-specific T cells *in vitro*. Because I-E⁺ islet cells are not completely resistant to immune attack *in vivo*, this also demonstrates, not surprisingly, that T cells under different conditions have varying susceptibilities to paralysis. T cells maintained *in vitro* appear to be unusually susceptible, but acutely activated T cells *in vivo* may be resistant to subsequent paralysis. T cells newly emigrated from the thymus may be most susceptible to paralysis, which could explain the observed tolerance in T cells from unmanipulated Ins-I-E mice¹⁰.

The findings reported here suggest a mechanism for peripheral tolerance induction to self tissue-specific antigens. It could be imagined that conditions such as infection, accompanied by

inflammation and local cytokine release, would induce class II expression by surrounding parenchymal (nonlymphoid) cells. Potentially autoreactive T cells that first encounter the combination of self tissue-specific antigen plus self class II on these cells rather than on M ϕ /DC would be paralysed, thus preventing the initiation of a 'bystander' autoimmune reaction. This might serve as an important fail-safe to complement the primary tolerance inducing function of the thymus. In some cases, other factors may tip the balance towards T-cell activation, and autoimmune responses would occur despite the expression of class II on parenchymal tissues.

The authors thank Dr J. Tomaszewski for assistance in evaluation of histology sections, M. Avarbock and F. Oram for technical assistance, and Dr C. Barker for his support. These studies were supported by NIH grants to A.N., R.D.P., R.L.B., and E.H.-K., and an American Cancer Society grant to E.H.-K. J.M. is a fellow of the Medical Scientist Training Program at the University of Pennsylvania.

Note added in proof: Gaspari *et al.*²³ have recently reported similar *in vitro* results using class II⁺ keratinocytes.

Received 17 August; accepted 17 October 1988.

- Heber-Katz, E., Hansburg, D. & Schwartz, R. H. *J. molec. cell. Immun.* **1**, 3-14 (1983).
- Buus, S., Sette, A., Colon, S. M., Jenis, D. M. & Grey, H. M. *Cell* **47**, 1071-1077 (1986).
- Bjorkman, P. J. *et al. Nature* **329**, 506-512 (1987).
- Frohman, M. & Cowing, C. *J. Immun.* **134**, 2269-2275 (1985).
- Lanzavecchia, A., Parodi, B. & Celada, F. *Eur. J. Immun.* **13**, 733-738 (1983).
- Pober, J. S. *et al. Nature* **305**, 726-729 (1983).
- Pujol-Borrell, R. *et al. Nature* **326**, 304-306 (1987).
- Bottazzo, G. F., Pujol-Borrell, R., Hanafusa, T. & Feldmann, M. *Lancet* **2**, 1115-1119 (1983).
- Bottazzo, G. F., Todd, I., Mirakian, R., Belfiore, A. & Pujol-Borrell, R. *Immun. Rev.* **94**, 137-169 (1986).
- Lo, D. *et al. Cell* **53**, 159-168 (1988).
- Sarvetnick, N., Liggitt, D., Pitts, S. L., Hansen, S. E. & Stewart, T. A. *Cell* **52**, 773-782 (1988).
- Allison, J. *et al. Nature* **333**, 529-533 (1988).
- Lafferty, K. J., Prowse, S. J., Simeonovic, C. J. & Warren, H. S. *A. Rev. Immun.* **1**, 143-173 (1983).
- Lacy, P. E. & Davie, J. M. *A. Rev. Immun.* **2**, 183-198 (1984).
- Heber-Katz, E., Valentine, S., Dietzschold, B. & Burns-Purzycki, C. *J. exp. Med.* **167**, 275-287 (1988).
- Quill, H. & Schwartz, R. H. *J. Immun.* **138**, 3704-3712 (1987).
- Jenkins, M. K. & Schwartz, R. H. *J. exp. Med.* **165**, 302-319 (1987).
- Happ, M. P., Kiraly, A. S., Offner, H., Vandenbark, A. & Heber-Katz, E. *J. Neuroimmun.* **19**, 191-204 (1988).
- Donohoe, J. A. *et al. Transplantation* **35**, 62-67 (1983).
- Bowen, K. M., Prowse, S. J. & Lafferty, K. J. *Science* **213**, 1261-1262 (1981).
- Brinster, R. L., Chen, H. Y., Trumbauer, M. E., Yagle, M. K. & Palmiter, R. D. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4438-4442 (1985).
- Lacy, P. E. & Kostianovsky, M. *Diabetes* **16**, 35-39 (1967).
- Gaspari, A. A., Jenkins, M. K. & Katz, S. I. *J. Immun.* **141**, 2216-2220 (1988).

Intestinal intraepithelial lymphocytes are a distinct set of $\gamma\delta$ T cells

Marc Bonneville, Charles A. Janeway Jr*, Kouichi Ito, Wayne Haser, Isao Ishida, Nobuki Nakanishi† & Susumu Tonegawa

Howard Hughes Medical Institute, and Center for Cancer Research, and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Lymphocytes are most reliably subdivided on the basis of their receptors for antigen at the cell surface. Three subtypes of lymphocytes are well defined: B cells that bear surface immunoglobulin and make antibody, CD4⁺T cells with CD3 $\alpha\beta$ receptors specific for antigen associated with class II major histocompatibility complex molecules, and CD8⁺T cells with CD3 $\alpha\beta$ receptors specific for antigen associated with class I MHC molecules. These T cells are responsible for known forms of cell-mediated immunity. The discovery of a third rearranging T-cell specific gene called γ (refs

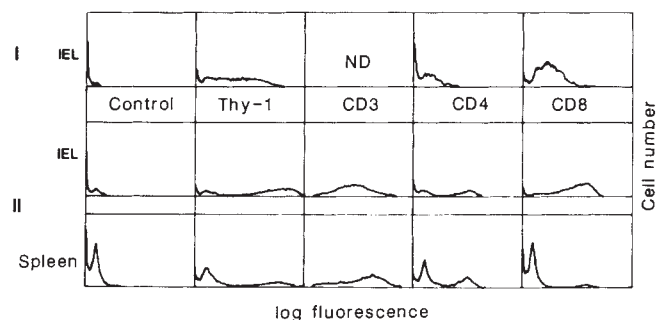


Fig. 1 FACS analysis of IELs and spleen cells. Two identical isolations were performed. IEL were prepared essentially as described¹². Cells were stained with biotinylated monoclonal antibodies to the structures shown, counterstained with FITC-avidin, and analysed on an EPICS flow cytometer. The FACS profiles show that most IELs are T cells expressing CD8.

1 and 2) has revealed the presence of a new class of T cells bearing a new receptor type, CD3 $\gamma\delta$ (refs 3-7). To date, neither the function nor the specificity of cells bearing this receptor has been determined. Because $\gamma\delta$ T cells are the main lymphocyte of epidermis^{8,9}, it was proposed that such cells could be important in surveillance of all epithelia¹⁰. We have isolated intraepithelial lymphocytes from murine small intestine¹¹⁻¹³, and shown that they predominantly or exclusively express CD3 $\gamma\delta$ receptors. Unlike the epidermal lymphocytes, these cells also express CD8, and they use a different V_γ gene to form their receptor. This strongly suggests that $\gamma\delta$ T cells home in a very specific manner to epithelia, where they presumably mediate their function.

T cells with $\alpha\beta$ receptors recognize fragments of foreign antigens associated with MHC molecules¹⁴, and can be subdivided on the basis of expression of CD4 for class II MHC recognition and CD8 for class I MHC recognition¹⁵. T cells with $\gamma\delta$ receptors have no known ligand and, in mouse, have been found to express neither CD4 nor CD8, being referred to as 'double negative' (DN) cells⁴⁻⁷. When cells expressing the murine pan T-cell marker Thy-1 were isolated and cloned from epidermis^{16,17}, these cells were shown to be DN CD3 $\gamma\delta$ T cells. It has recently been shown that such cells almost exclusively express the products of the $V_{\gamma 5}$ gene (see refs 6 and 26 for the nomenclature of mouse V_γ gene segments) associated with $V_{\delta 1}$ (ref. 9; W. L. Havran and J. P. Allison, and J. Levis, P. Tucker, P. Bergstresser and R. Tigelaar, personal communication). Such cells are also the earliest T cells to appear in the thymus at about day 15 of gestation, and they are not found in the thymus after birth¹⁸. The homogeneity of these dendritic epidermal cells (DECs) suggests a very specific homing mechanism.

It is well known that many epithelia contain lymphoid populations interdigitating between the basolateral faces of epithelial cells. The best characterized of these are the intestinal intraepithelial lymphocytes (IELs), which have been isolated in several species. These cells are Thy-1⁺, CD4⁻, and mainly CD8⁺ (refs 11-13). They have been shown to have cytolytic activity in a number of non-specific cytotoxic assays^{11,12}. The nature of their receptor or their ligand specificity is not known. If such cells were shown to be mainly or exclusively $\gamma\delta$ T cells, this would extend the evidence that $\gamma\delta$ T cells home to epithelia, and would imply that such $\gamma\delta$ lymphocytes could be involved in surveillance of epithelial cell integrity; they could thus represent a first line of defence against infection and transformation of epithelial cells.

To determine the nature of the receptor on IELs, we isolated such cells from murine small intestines^{13,19}. As previously shown¹¹⁻¹³, these cells were mainly Thy-1⁺, CD8⁺, CD4⁻ (Fig. 1, preparation I); in some preparations, a significant number of CD4⁺T cells were found (Fig. 1, preparation II). These probably represent contaminants from lamina propria or from Peyer's patch, but the techniques used in these experiments cannot

* Permanent address: Howard Hughes Medical Institute, Department of Pathology and Section of Immunobiology, Yale University School of Medicine, New Haven, Connecticut 06510, USA.

† Present address: Howard Hughes Medical Institute, Columbia University, New York, New York 10032, USA.

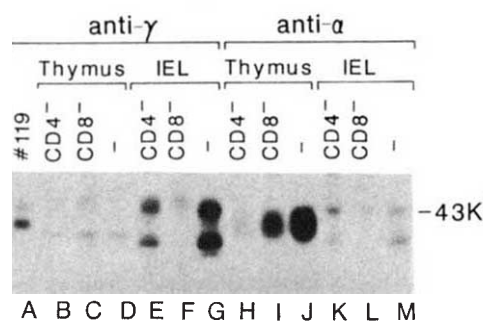


Fig. 2 Immunoprecipitated receptors on IEL and thymocytes. IEL and thymocytes were untreated, or treated with anti-CD4 or anti-CD8 and complement, the surviving cells labelled with ^{125}I , and lysates precipitated with anti- γ (lanes A–G) or subsequently with anti- α (lanes H–M), and resolved on SDS-PAGE minigels. Material in lanes B–D, H–J is derived from thymus; lanes E–G, K–M, from IEL; lane A, from hybridoma 119 expressing $V_{\gamma 1}$. Material in lanes B, E, H and K has been treated with anti-CD4; lanes C, F, I and L, treated with anti-CD8 and complement. IELs express large amounts of $\gamma\delta$ receptor, depleted by anti-CD8 and complement, whereas thymus has less $\gamma\delta$ receptor not depleted by anti-CD8. IELs lack $\alpha\beta$ receptors; bands precipitated represent residual $\gamma\delta$ receptor from the first immunoprecipitation. Note depletion of $\alpha\beta$ receptors in thymocytes by anti-CD4 and anti-CD8.

Methods. IEL and thymocytes were isolated from adult (nine weeks old) C57BL/6 mice, divided into aliquots of 2×10^6 IELs or 2×10^7 thymocytes, treated for 60 min at 37°C with anti-CD4 (RL 172.4) or anti-CD8 (3.155.D14) plus rabbit complement (Lowtox, Cedarlane); live cells were recovered by centrifugation over Ficoll-hypaque, the six pools of cells were labelled with ^{125}I and lysed with NP40 in buffer⁶. Lysates were precipitated with anti- γ monoclonal antibody, immune complexes and beads were centrifuged out, and the remaining supernatant was subjected to a second immunoprecipitation with beads coated with purified anti- α antiserum. Anti- α antiserum was prepared in rabbits by immunization with a fusion protein of β -galactosidase constructed by cloning a 655-base pair fragment of the α -chain gene of pHSD58 isolated by digestion with *RSAI* and *HincII* into $\lambda\text{gt}11$. The clone was expressed in *Escherichia coli* and the fusion protein was purified by affinity chromatography on a column of anti- β -galactosidase. Rabbits were hyper-immunized with this protein, and the immunoglobulin component purified by protein A Sepharose chromatography. Anti- γ monoclonal antibody KN365 was raised by immunization of DBA/2 mice with a bacterial recombinant-derived γ -chain fragment consisting of 20 amino acids of $V_{\gamma 2}$, all of $J_{\gamma 2}$, and 123 amino acids of $C_{\gamma 2}$. KN365 reacts with an epitope on $C_{\gamma 1}$ and $C_{\gamma 2}$ regions after denaturation of the protein with SDS (ref. 23).

determine this conclusively. These isolated cells were then labelled with ^{125}I , and the receptors immunoprecipitated using either anti- γ monoclonal antibody or a purified anti- α antiserum. As can be seen in Fig. 2, there are abundant $\gamma\delta$ heterodimers on IELs (Fig. 2, lane G), but little $\gamma\delta$ detectable on adult thymocytes from the same donors (Fig. 2, lane D). Thymocytes show abundant $\alpha\beta$ receptors, whereas IELs show no detectable $\alpha\beta$ receptor (Fig. 2, lanes J and M respectively).

As virtually all the IELs isolated express CD8, and as most or all of the receptor protein precipitated is $\gamma\delta$, it seems likely that the T cells expressing $\gamma\delta$ receptors are mainly CD8⁺. We and others have previously shown that anti-CD8 plus complement treatment of thymocytes kills most of the cells, but does not deplete cells expressing $\gamma\delta$ receptors^{4–7}; we have confirmed this result, see Fig. 2, lanes C and D. Using IELs, however, anti-CD8 plus complement almost totally eliminates T cells bearing $\gamma\delta$ receptors (Fig. 2, lanes F and G). Thus, most (or all) of the CD8⁺ T cells isolated from the gut express $\gamma\delta$ T-cell receptors. As a control, anti-CD4 plus complement was shown not to deplete $\gamma\delta$ T cells from either thymocytes (Fig. 2, lane B) or IEL (Fig. 2, lane E), as expected. Finally, sequential

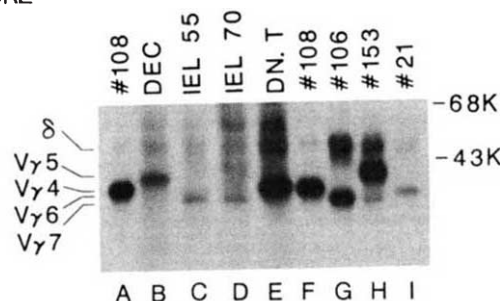


Fig. 3 Immunoprecipitation of the $\gamma\delta$ protein complex from DEC, IELs, DN thymocytes and T-hybridomas. DEC, IELs, DN thymocytes and selected T-cell hybrids were surface iodinated, and $\gamma\delta$ proteins were precipitated with anti- γ monoclonal antibody and resolved on one-dimensional gel electrophoresis. Lane B, DEC; lane C, IEL from 55% Percoll gradient; lane D, IEL from 70% Percoll gradient; lane E, DN thymocytes; lanes A, F, G, H and J, $\gamma\delta$ -expressing hybridomas (number 108: $V_{\gamma 4}V_{\delta 1}$; number 106: $V_{\gamma 5}V_{\delta 1}$; number 153: $V_{\gamma 5}V_{\delta 1}$ and number 21: $V_{\gamma 6}V_{\delta 1}$). The reference mobilities of γ -chains from hybridomas using $V_{\gamma 4}$, $V_{\gamma 5}$, $V_{\gamma 6}$ and $V_{\gamma 7}$ are indicated on the left. Note the presence of three species immunoprecipitated from $V_{\gamma 5}$ -expressing hybridoma number 153. The lowest species corresponds to a deglycosylated $V_{\gamma 5}J_{\gamma 1}C_{\gamma 1}$ product⁹.

Methods. DEC were prepared from trypsinized epidermal sheets as described²⁴ and incubated overnight in culture medium at 37°C before immunoprecipitation to allow recovery of their surface receptors. IELs were isolated according to the technique described by Petit *et al.*¹⁹ and lymphocytes from 55% and 70% Percoll gradients were analysed separately. DN thymocytes were obtained after complement-dependent lysis of CD4⁺ and CD8⁺ cells, as described in Fig. 2. T-cell hybrids were prepared by fusing DN thymocytes from mice at different ages with BW5147 thymoma, selected in HAT medium, and screened, cloned, and characterized as described elsewhere^{18,20}. Cells were iodinated and lysates immunoprecipitated with anti- γ monoclonal antibody KN365 as described previously^{5,6}.

precipitation of the same lysates with anti- α antibody precipitates no $\alpha\beta$ receptor proteins from IELs and demonstrates that anti-CD8 plus complement, and especially anti-CD4 plus complement, greatly deplete $\alpha\beta$ receptor in the thymocyte population (Fig. 2, lanes H–J).

In separate experiments, we have identified the $\gamma\delta$ T-cell receptor proteins of a series of T-cell hybrids derived from fetal and adult DN thymocytes fused with the T-cell lymphoma Bw5147 (refs 18 and 20). By analysing DNA and RNA, we could determine which γ - and δ -genes encode the $\gamma\delta$ heterodimers expressed on each of these hybridomas^{18–20}. This analysis revealed that γ -chains encoded by different V_{γ} gene segments migrate with distinct mobilities in SDS-PAGE, as shown in Fig. 3, lanes A, F, G, H and I. On the basis of this criterion the IEL γ -chain is clearly distinct from those encoded by $V_{\gamma 4}$ and $V_{\gamma 5}$ which are the chief γ -chain types expressed on adult thymocytes (ref. 20 and Fig. 3, lane E) and DEC (ref. 9; lane B) respectively. Instead, IEL γ chains seem to comigrate, either with $V_{\gamma 7}$ - or $V_{\gamma 6}$ -encoded γ -chains (Fig. 3, lanes C, D, G and I).

In a further attempt to assign the IEL γ -chains to a specific V_{γ} gene, we examined RNA present in IEL using V_{γ} -specific hybridization probes. As shown in Fig. 4, $V_{\gamma 7}$ probe detected a moderate level of γ RNA in IELs (Fig. 4a, lanes F and G), whereas RNA was absent or present at a barely detectable level when $V_{\gamma 4}$ (Fig. 4b, lanes F–G), $V_{\gamma 6}$ (Fig. 4c, lanes F–G) or $V_{\gamma 5}$ (data not shown) probes were used. This contrasted with the results from DN thymocytes, where we found large amounts of $V_{\gamma 4}$ transcripts (Fig. 4b, lane E) and very low or zero levels of $V_{\gamma 7}$, $V_{\gamma 6}$, or $V_{\gamma 5}$ mRNAs (Fig. 4a and c, lane E, and data not shown). These results corroborate those for the proteins and indicate that IEL γ chain could be primarily encoded by the $V_{\gamma 7}$ gene segment.

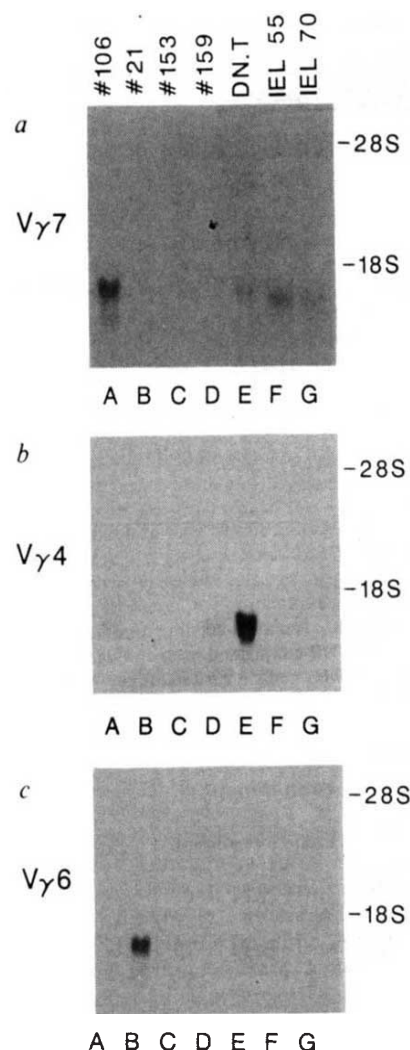


Fig. 4 Northern blot analysis of IELs and DN thymocytes. Total RNA from IELs, DN thymocytes and T-cell hybrids were electrophoresed in agarose gel, blotted to nylon membranes and hybridized to $V_{\gamma 7}$ (a), $V_{\gamma 4}$ (b) or $V_{\gamma 6}$ (c) probes. Lanes A–D: T-cell hybridomas (number 106: $V_{\gamma 7}$; number 21: $V_{\gamma 6}$; number 153: $V_{\gamma 5}$; number 159: $V_{\gamma 4}$); lane E: DN thymocytes; lane F: IEL from 55% Percoll gradient; lane G: IEL from 70% Percoll gradient. Hybridoma number 106 expresses both $V_{\gamma 4}$ and $V_{\gamma 7}$ mRNAs but analysis of cDNAs revealed that the $V_{\gamma 4}J_{\gamma 1}C_{\gamma 1}$ gene was joined out-of-frame²⁰. **Methods.** IELs, DN thymocytes and T-cell hybrids were prepared as described in Fig. 3. Total RNA (5 μ g) was electrophoresed, blotted on to nylon membranes, prehybridized and hybridized with random primed probes^{18,20}. Details of $V_{\gamma 4}$, $V_{\gamma 5}$, $V_{\gamma 6}$ and $V_{\gamma 7}$ probes are described elsewhere^{18,20}.

Our data have shown that another significant intraepithelial lymphocyte population besides DEC, that of the small intestine, is made up of T cells with $\gamma\delta$ receptors. It should be noted that our analysis applies to the cells isolated by two different procedures described in Fig. 1 and Fig. 3. Nevertheless, to establish whether these cells are representative of all IELs will require an *in situ* approach. Assuming the isolated IEL population is representative, then IELs differ from DEC in two important respects: IELs express CD8 and IELs use a different $V_{\gamma}V_{\delta}$ combination to form receptors. These data have several important consequences.

First, that both epidermis and intestinal epithelium are sites of predominant $\gamma\delta$ T-cell populations suggests that such cells have some critical influence in epithelia. Indeed, the virtual absence of T cells expressing $\alpha\beta$ receptors in these tissues, and the fact that the unusual CD8⁺ $V_{\gamma 7}V_{\delta 6}$ phenotype of the IEL is seldom found in any somatic lymphoid organ²⁰, both suggest that such cells are specialized for homing to intestinal epithelium. That IELs and DEC differ in the V regions used in forming their receptors further points to a high degree of tissue specificity in this homing pattern. This, in turn, strongly suggests that at least these populations of $\gamma\delta$ T cells are specialized for epithelial surveillance in particular sites, as discussed²¹.

Although it has been difficult to demonstrate that DEC are present in species other than the mouse, it is clear that there are IELs in chickens and man, and that these cells also express $\gamma\delta$ receptors and CD8 (R. P. Bucy *et al.*, personal communication). The restricted V-gene usage in these epithelial sites suggests that ligand recognition in these tissues is limited. Whether such receptors are truly autoreactive to MHC class¹⁰, or whether

they have some different function and specificity altogether, remains to be determined. Likewise, the mechanism of homing or selective retention of such cells in particular epithelial sites is not understood. Nevertheless, the finding of different $\gamma\delta$ T cells in different epithelia strongly points to a highly specific biological function for such cells in these anatomical locations.

After this study was completed, we became aware that T. Goodman and L. Lefrançois²² independently showed that CD8⁺ T cells isolated from intestinal epithelium express $\gamma\delta$ receptors, although they do not specify which γ -gene is used in this site.

We thank Drs P. B. Ernst and J. Bienenstock for help in preparing some of the IEL populations. C.A.J. thanks the Howard Hughes Medical Institute for sponsorship of a sabbatical leave. The authors thank Elly Basel for invaluable secretarial assistance. We acknowledge the contribution of Drs Kenya Shitara and Hajime Yoshida of Kyowa Hakko Kogyo Co Ltd in the preparation of the anti- $\gamma\delta$ monoclonal antibody. This work was supported by grants from the Howard Hughes Medical Institute, American Cancer Society, Ajinomoto Co. Ltd and the NIH.

Received 8 June; accepted 21 October 1988.

1. Saito, H. *et al.* *Nature* **309**, 757–762 (1984).
2. Hayday, A. C. *et al.* *Cell* **40**, 259–269 (1985).
3. Brenner, M. B. *et al.* *Nature* **322**, 145–149 (1986).
4. Bank, L. *et al.* *Nature* **322**, 179–81 (1986).
5. Nakanishi, N., Maeda, K., Ito, K., Heller, M. & Tonegawa, S. *Nature* **325**, 720–723 (1987).
6. Maeda, K. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **84**, 6536–6540 (1987).
7. Lew, A. M. *et al.* *Science* **234**, 1401–1405 (1986).
8. Koning, F. *et al.* *Science* **236**, 834–837 (1987).
9. Kuziel, W. A. *et al.* *Nature* **328**, 263–266 (1987).
10. Janeway, C. A. Jr, Jones, B. & Hayday, A. C. *Immun. Today* **9**, 73–76 (1988).
11. Ernst, P. B., Befus, A. D. & Bienenstock, J. *Immun. Today* **6**, 50–55 (1985).
12. Klein, J. R. *J. exp. Med.* **164**, 309–313 (1986).
13. Davies, M. D. J. & Parrott, D. M. *V. Gut* **22**, 481–490 (1981).
14. Germain, R. N. *Nature* **322**, 687–689 (1986).
15. Swain, S. L. *Immunol. Rev.* **74**, 129–156 (1983).
16. Stingl, G. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **84**, 2430–2434 (1987).
17. Takashima, A., Nixon-Fulton, J. E., Bergstresser, P. R. & Tiegelaar, R. E. *J. invest. Derm.* **90**, 671–678 (1988).
18. Ito, K. *et al.* *Proc. natn. Acad. Sci. U.S.A.* (submitted).
19. Petit, A. *Eur. J. Immun.* **15**, 211–215 (1985).
20. Takagaki, Y. *et al.* *J. Immun.* (submitted).
21. Janeway, C. A. Jr *Nature* **333**, 804–806 (1988).
22. Goodman, T. & Lefrançois, L. *Nature* **333**, 855–858 (1988).
23. Nakanishi, N. thesis, Univ. Tokyo (1987).
24. Nixon-Fulton, J. L., Bergstresser, P. R. & Tigelaar, R. E. *J. Immun.* **136**, 2276–2286 (1986).
25. Matis, L. A., Cron, R. & Bluestone, J. A. *Nature* **330**, 262–264 (1987).
26. Hellig, J. S. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **84**, 8070–8074 (1987).

Killing of antigen-reactive B cells by class II-restricted, soluble antigen-specific CD8⁺ cytolytic T lymphocytes

Nobukata Shinohara*†, Motoo Watanabe‡, David H. Sachs* & Nobumichi Hozumi‡

* Immunology Branch, National Cancer Institute, NIH Bethesda, Maryland 20892, USA

‡ Mount Sinai Hospital, Research Institute, Toronto, Ontario, Canada

Cytolytic T lymphocytes (CTLs) are generally thought to recognize cellular antigens presented by class I MHC molecules. A number of studies, however, have revealed responses of considerable magnitude involving both CD8⁺ and CD4⁺ CTLs with class II restriction^{1–5}, suggesting that class II-restricted CTLs recognizing exogenous protein antigens may exist. As class II antigens are normally expressed on limited types of cells such as B cells and macrophages, such CTLs might be expected to exert a suppressive

† Present address: Laboratory of Cellular Immunology, Mitsubishi Kasei Institute of Life Sciences, Tokyo, Japan.

Fig. 1 Specific lysis of B-cell targets by T-cell lines induced with soluble antigen (Ag). Lymphocytes were obtained from inguinal and paraaortic lymph nodes of BALB/c mice immunized by injection of 100 µg of KLH or OVA in complete Freund's adjuvant (CFA) into the base of the tail 7–10 days before collection. The cell lines were depleted of either CD8⁺ (CD4 BALB@OVA line) or CD4⁺ (CD8 BALB@KLH line) cells by treating with the appropriate monoclonal antibody plus complement, and stimulated *in vitro* with antigen in the presence of 10% TCGF (48-h culture supernatant of Con A-stimulated rat lymphocytes). After the second stimulation, irradiated (2,000 R) B10.D2 spleen cells were added as antigen-presenting cells. Cells were tested for specific lytic activities on A20.HL, a class II-antigen-positive BALB/c (H-2^d) B-cell line transfected with structural genes of heavy and light chains of anti-TNP IgM antibody⁶, in the presence of 10 µg ml⁻¹ of TNP-conjugated or unconjugated antigens. B10.A@B10.D2 line was established by repeatedly stimulating B10.A lymphocytes with irradiated B10.D2 cells *in vitro* in the presence of 5% TCGF. This line was used as a positive control for assessing lysability of the target cells. Effector: target (E:T) ratio was 8:1 for each group. @ stands for anti-.

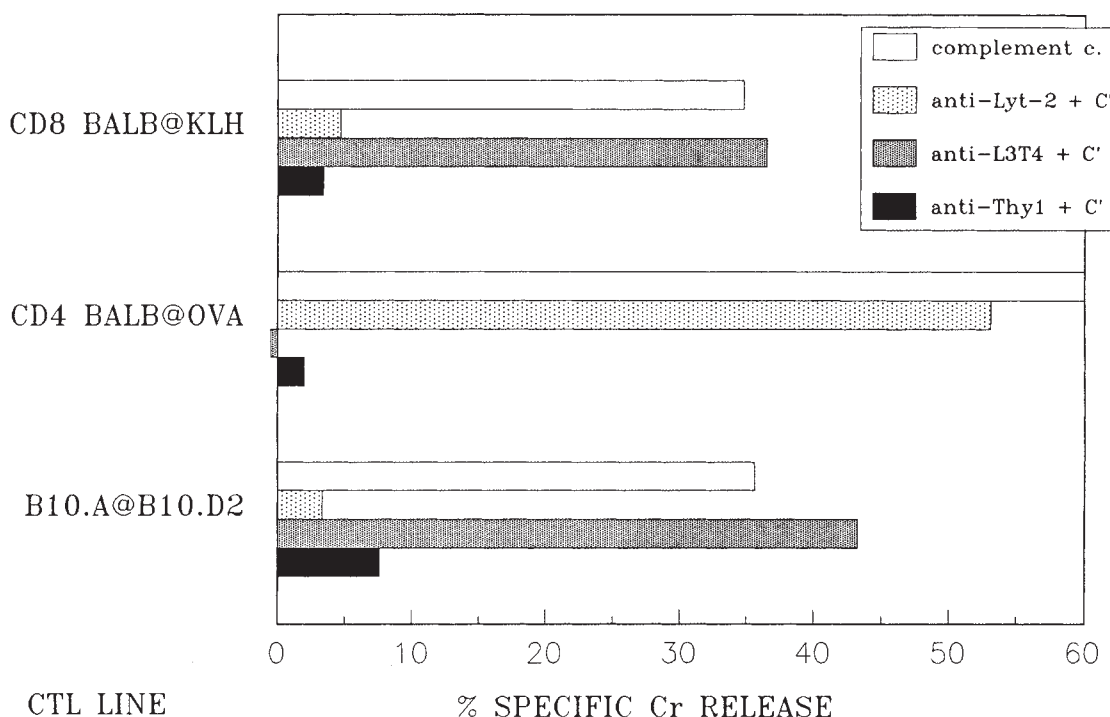
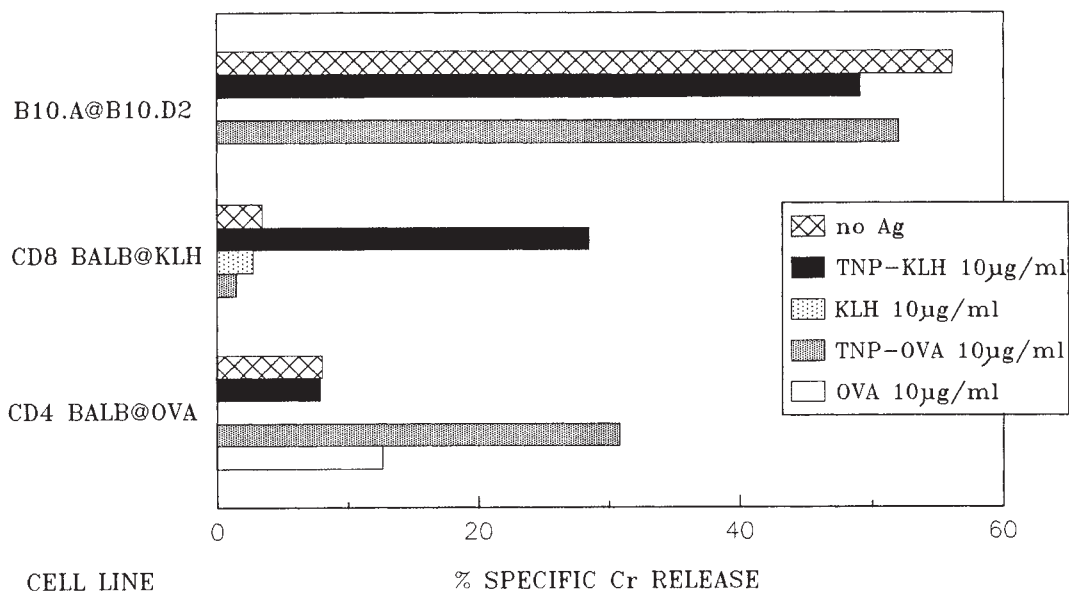


Fig. 2 Phenotypes of soluble antigen-specific CTL lines. CTL lines were tested on A20.HL cells pre-incubated overnight with TNP-KLH (for CD8⁺ BALB@KLH line) or TNP-OVA (for CD4⁺ BALB@OVA and B10.A@B10.D2 lines). The washed target cells were labelled with ⁵¹Cr. CTL lines plated in 96-well U-bottomed plates were treated with monoclonal antibodies to CD8 (83-12-5) or CD4 (GK1.5) plus complement. After washing the treated cells in the plate, cells were resuspended in 100 µl of medium, labelled target cells were added to the wells, and six-hour CML assays were performed. The E:T ratio was 8:1 for each group.

effect on antibody responses. Here we report that stimulation of mouse lymphocytes with a soluble antigen induced CD8⁺ and CD4⁺ CTLs specific for the antigen with class II restriction. The specific lysis was far more efficient when target B cells specifically recognized the antigen than when they did not, indicating that the primary targets for these CTLs are probably B cells expressing immunoglobulin receptors reactive for the same antigen molecule. These results suggest that the natural occurrence of such CTLs during immune responses may explain antigen-specific suppression on antibody responses by T cells.

Lymph node lymphocytes obtained from BALB/c mice immunized with keyhole limpet haemocyanin (KLH) were depleted of CD4⁺ cells by antibody plus complement and were stimulated *in vitro* several times with antigen in the presence of 10% T cell growth factor (TCGF) to establish a bulk line, CD8⁺BALB anti-KLH. A control bulk line, CD4⁺BALB anti-ovalbumin (OVA) was similarly developed. These cells were tested for specific lytic activities on A20.HL, a class II-antigen-positive BALB/c (H-2^d) B-cell line transfected with heavy and light chain structural genes of an anti-TNP (2,4,6-trinitrophenyl)

Fig. 3 Dose effect of antigen concentration on specific lysis. ^{51}Cr -labelled A20.HL or A20.2J cells plated in 96-well U-bottomed plates were incubated in the presence of graded concentrations of antigens for two hours. CTL lines were added after the incubation and six-hour CML assays were performed. E:T ratio was 8:1.

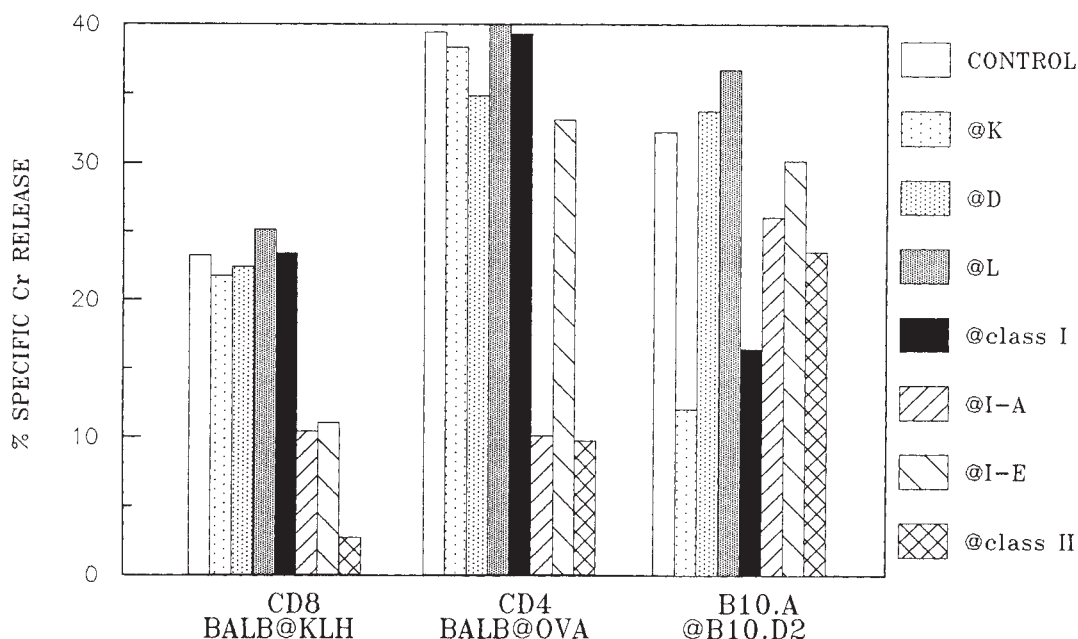
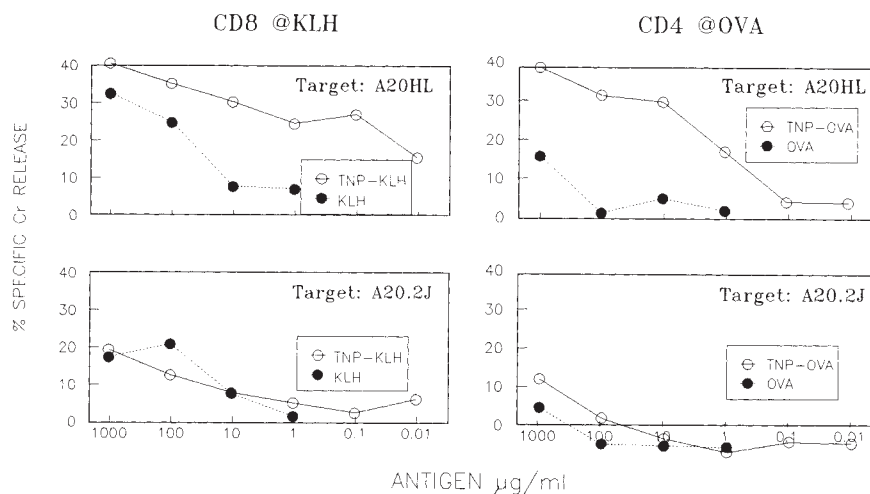


Fig. 4 Sensitivity of specific lysis to blocking by monoclonal antibodies to MHC antigens. $10\ \mu\text{l}$ of 1:10 dilution of the ascitic form of monoclonal antibodies were added to CML assay cultures. Antibodies used were reactive with H-2K^d (20-8-4), D^d (15-1-5), L^d (30-5-7), I-A^d (25-9-17), and I-E^d (14-4-4). @class I and @class II indicate mixtures of antibodies to class I (K, D, L) and class II (I-A and I-E), respectively. E:T ratio was 8:1.

IgM antibody⁶, in the presence of $10\ \mu\text{g ml}^{-1}$ of TNP-conjugated and unconjugated antigens (Fig. 1). The CD8⁺BALB anti-KLH line significantly lysed the transfectant in the presence of TNP-KLH but not KLH in a seven-hour cell-mediated lympholysis (CML) assay. The antigen-specific nature of the killing was corroborated, as the OVA-induced CD4⁺ line had reciprocal specificity of lysis. Phenotypes of the cells responsible for the observed antigen-specific lysis were confirmed by treatment of the effector cells with anti-CD8 or anti-CD4 antibodies and complement (Fig. 2). The lytic activity of the CD8⁺ KLH-specific line was totally abrogated by anti-CD8 treatment, whereas specific lysis by the CD4⁺ OVA-specific line was sensitive to anti-CD4. These results therefore demonstrated that there are both CD8⁺ and CD4⁺ CTL with specificity for soluble antigens.

The concentration of antigen required to sensitize A20.HL cells for specific lysis was determined by pre-incubating A20.HL targets for two hours in the presence of graded concentrations of antigen before adding CTLs for a further six-hour incubation (Fig. 3). Specific lysis of the anti-TNP transfected A20.HL by the CD8⁺ anti-KLH line was observed when $\geq 100\ \text{ng ml}^{-1}$ of TNP-KLH was present in CML culture whereas much higher concentrations ($\geq 100\ \mu\text{g ml}^{-1}$) were required for native KLH

to induce a comparable level of killing. When the non-transfected parental cell (A20.2J) was used as target, specific lysis was detectable only at high antigen concentrations and no significant difference was seen between TNP-conjugated and unconjugated KLH. Similar results were obtained with the CD4⁺ anti-OVA line. These results indicate that specific killing of the B-cell target is highly dependent on the ability of the B cells to bind the antigen molecule via the immunoglobulin receptor. As it is unlikely that such high antigen concentration ($\geq 100\ \mu\text{g ml}^{-1}$) could be attained *in vivo* during an ordinary immunization procedure, the primary targets for these CTLs are probably the B cells that are reactive with the same molecule.

CML assays were carried out on A20.HL pre-pulsed with the relevant antigen in the presence of a panel of monoclonal antibodies directed to products of the five known MHC loci of the H-2^d haplotype (K, D, L, I-A and I-E) (Fig. 4). Specific target-cell lysis by the CD8⁺ anti-KLH line was moderately inhibited by anti-I-A and anti-I-E antibodies. Lysis was almost completely inhibited when the two anti-class II antibodies were added together, suggesting that the line contained a mixture of I-A- and I-E-restricted clones. Specific lysis by the CD4⁺ OVA-specific line was drastically reduced in the presence of the

anti-I-A antibody alone. Neither of the two antigen-specific lines were sensitive to blocking by anti-class I antibodies.

This report demonstrates that immunization with exogenous protein antigens induces both CD8⁺ and CD4⁺ CTLs specific for the antigen presented in the context of class II MHC antigens. Specific lytic activity of CD4⁺ helper T cells has previously been documented by several investigators⁷⁻¹¹, and the antigen specific lysis of A20.HL by the CD4⁺ BALB anti-OVA line reported here probably represents another example of such lytic activity. As far as we know, however, the existence of CD8⁺ CTLs specific for soluble antigens presented by class II molecules has not previously been reported. The CD8⁺ soluble antigen-specific CTLs recently reported by Staerz *et al.*¹² were clearly different from those reported here, as they were class I-restricted whereas the vast majority of CD8⁺ CTLs detected in our experiments showed class II-restriction. Differences in the assay system between the two laboratories may explain this discrepancy as we have used class II-positive B-cell targets bearing receptors specific for the antigen, whereas Staerz *et al.* used class II-negative, surface-immunoglobulin-negative, target cells. In our studies, presentation of the antigen on target cells probably involved specific binding and antigen processing by these target cells, because native antigen at low concentration was used in the effector phase. In the studies by Staerz *et al.*, high concentrations of peptide fragments were necessary to reveal specific lysis¹². Such conditions probably allowed antigen fragments to bind directly to class I molecules on the surface of target cells without specific binding or intracellular processing, as has been demonstrated for influenza-specific class I-restricted CTL¹³. Thus it seems likely that the two assay systems would selectively detect different types of CTLs.

The physiological contribution of CTLs specific for soluble

antigens to immune responses is not entirely clear. Nevertheless, these observations, together with the reports on antigen-specific lysis by CD4⁺ cells⁷⁻¹¹, suggest a new possible interpretation of specific suppression of immune responses. The class II-restricted CTLs specific for soluble antigens reported here might be involved in specific lysis of antigen-binding B cells in the presence of physiological concentrations of native antigen. Such CTLs might provide an explanation of at least a fraction of the antigen-specific suppression of humoral antibody responses by T cells¹⁴. One major argument against this explanation would be that both CD4⁺ and CD8⁺ populations were found to cause specific lysis, whereas specific suppression has generally been attributed to the function of CD8⁺ T-cell populations¹⁴. However, potent helper activity within the CD4⁺ sub-populations might have obscured the suppressive effects of CD8⁺ CTLs. In any case, these results suggest that a re-examination of the mechanism of antigen-specific suppression of humoral responses by T cells may be worthwhile.

Received 11 August; accepted 3 October 1988.

1. Vidovic, D., Juretic, A., Nagy, Z. A. & Klein, J. *Eur. J. Immun.* **1**, 499-509 (1981).
2. Shinohara, N. & Kojima, M. *J. Immun.* **132**, 578-583 (1984).
3. Shinohara, N. *et al. J. Immun.* **140**, 30-36 (1988).
4. Golding, H. & Singer, A. *J. Immun.* **135**, 1610-1615 (1985).
5. Kaplan, D. R., Griffith, R., Braciale, V. L. & Braciale, T. J. *Cell. Immun.* **88**, 193 (1984).
6. Watanabe, M., Wegmann, D. R., Ochi, A. & Hozumi, N. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5247-5251 (1986).
7. Tite, J. P., Powe, M. B. & Ruddle, N. H. *J. Immun.* **135**, 25-33 (1985).
8. Watanabe, M., Yoshikawa, M. & Hozumi, N. *Immun. Lett.* **15**, 133-138 (1987).
9. Ozaki, S., York-Jolley, J., Kawamura, H. & Berzofsky, J. A. *Cell. Immun.* **105**, 301-316 (1987).
10. Nakamura, M., Ross, D. T., Briner, T. J. & Geffer, M. L. *J. Immun.* **136**, 44-47 (1986).
11. Tite, J. P. & Janeway, C. A. Jr *Eur. J. Immun.* **14**, 878-880 (1984).
12. Staerz, U. D., Karasuyama, H. & Garner, A. M. *Nature* **329**, 449-451 (1987).
13. Townsend, A. R. M. *et al. Cell* **44**, 959-968 (1986).
14. Tada, T., Taniguchi, M. & Takemori, T. *Immun. Rev.* **26**, 106-129 (1975).

HIV-1 gag-specific cytotoxic T lymphocytes defined with recombinant vaccinia virus and synthetic peptides

Douglas F. Nixon*, Alain R. M. Townsend*, John G. Elvin*, Charles R. Rizza†, John Gallwey‡ & Andrew J. McMichael*

* Institute of Molecular Medicine, John Radcliffe Hospital, Oxford OX3 9DU, UK

† Oxford Haemophilia Centre, Churchill Hospital, Oxford OX3 7LJ, UK

‡ Harrison Department, Radcliffe Infirmary, Oxford OX2 6HE, UK

Current candidate vaccines fail to protect primates against challenge with human immunodeficiency virus (HIV) in the presence of antibody responses¹; this underlines the importance of studying cell-mediated immunity to HIV and identifying specific epitopes that stimulate cytotoxic T lymphocytes (CTL). Using a recombinant vaccinia virus to express the gag protein of HIV-1 we found HLA class-I-restricted gag-specific CTL in thirteen out of fifteen healthy HIV seropositive patients. We then used short synthetic peptides in the lysis assay²⁻⁵ to screen for gag CTL epitopes. In one patient we have identified a peptide in p24 that is recognized by CTL in association with HLA-B27. This peptide, and further peptide sequences defined by these methods, could be incorporated in vaccines designed to induce cell-mediated immunity against HIV.

Previous studies have demonstrated CTL specific for the envelope (env) and reverse transcriptase (pol) proteins⁶⁻⁸ in most healthy HIV seropositive patients; a minority also had weak CTL responses directed against the group-specific antigen (gag)⁶. To study the gag-specific response we constructed a

recombinant vaccinia virus that expressed the gag protein of HIV-1 (gag-vac). Our initial studies showed that freshly prepared peripheral blood mononuclear cells (PMB) from four consecutive HIV seropositive donors did not lyse autologous B lymphoblastoid cells infected with the gag-vac. Therefore a small fraction of these PBM were activated with the mitogen phytohaemagglutinin (PHA) for 24 hours and then added back to the remaining lymphocytes that were cultured *in vitro*. A proportion of the activated CD4-positive cells should express HIV proteins under these conditions. As shown in Fig. 1a, effector cells restimulated in this way specifically lysed autologous lymphoblastoid target cells infected with gag-vac. At the highest effector: target (E:T) ratios some lysis of the control targets, uninfected and infected with another recombinant vaccinia virus, was observed with some individuals (see also Table 1). This may indicate some nonspecific mechanism, such as natural killer-cell-mediated lysis. In the experiments with HIV seropositive donors lysis of the gag-vac infected targets was clearly above this background lysis and was shown to be HLA restricted (Fig. 2). PBM cultured in the absence of PHA-activated T cells failed to respond. PBM from six seronegative volunteers failed to lyse gag-vac infected autologous cells. One of these results is shown in Fig. 1b. In the patients shown in Fig. 2, HLA-restricting elements were identified. B lymphoblastoid cells, that shared single HLA antigens with effector cells, were tested as target cells. The gag-specific CTL response of donor 007 was restricted through HLA-B27 (Fig. 2a) and HLA-A1 (Fig. 2b) but not B44. Whereas the B27-restricted response has been constant over a six-month period, the A1-restricted response was not always observed, which suggests that the clonal composition of the CTL response may vary over time. In another donor, 008, an HLA-A2 restricted response was seen (Fig. 2c); HLA-B18- and B14-restricted responses have also been observed. Thus, there are likely to be several epitopes recognized by CTL in the gag protein that are restricted by different HLA antigens.

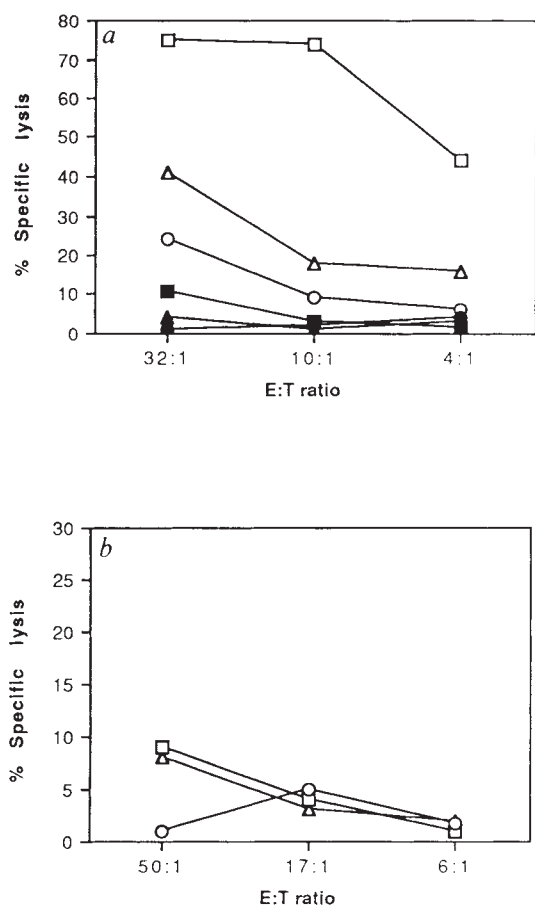


Fig. 1 *a*, Induction of HIV gag-specific CTL *in vitro*. Lysis by effector cells from a seropositive donor cultured *in vitro* with autologous PHA-activated lymphocytes (open symbols), or PBM cultured *in vitro* (closed symbols), on autologous lymphoblastoid cells infected with gag-vac (□, ■), or, as controls, M-vac (△, ▲) or uninfected (○, ●). *b*, Failure to induce gag-specific CTL in PBM of an HIV seronegative donor. Effector cells cultured *in vitro* with autologous PHA-activated lymphocytes tested on autologous B lymphoblastoid cells uninfected (○), infected with M-vac (△), or gag-vac (□).

Methods. Peripheral blood (40 ml) from donors were separated on Ficoll-Hypaque, lymphocytes washed and resuspended in RPMI-1640 (Gibco Biocult) with 10% fetal calf serum (Gibco Biocult). Autologous PHA-activated lymphocytes were prepared by culturing $5-10 \times 10^6$ PBM with PHA (Wellcome) at $2 \mu\text{g ml}^{-1}$ for 24 h at 37°C ; they were washed twice in RPMI-1640-10% FCS before adding to the remaining unstimulated PBM and were cultured together for seven days. Addition of T-cell growth factor (Lymphocult T, Biotest) at 10 units ml^{-1} seven days after addition of PHA-activated cells, resulted in growing T-cell lines that showed the same specificity and grew for up to four weeks. In the cytolytic assay target cells were Epstein-Barr virus transformed cell lines, infected (at 10 plaque forming units per cell) with recombinant vaccinia virus or without, and labelled with ^{51}Cr (Amersham). Specific lysis was measured in a 5-h ^{51}Cr release assay at the effector:target (E:T) ratios shown as previously described². Gag-vac was constructed as follows: Plasmid pIII 5.5 (ref. 11), a gift from Dr R. Gallo (NCI, Bethesda, Maryland, USA) was cut with the restriction enzymes *Bss*HI and *Hinc*II, the gag-gene fragment was isolated, its ends filled in and inserted into the plasmid pSC11 (ref. 12) and then into vaccinia virus¹³ under the 7.5 K early-late promoter. B lymphoblastoid cells infected with this virus expressed the p55 gag structural protein which was detectable by immunoprecipitation (data not shown). Recombinant vaccinia virus that expresses influenza A virus matrix protein (M-vac) was a gift from Drs G. Smith and B. Moss (NIH, Bethesda, Maryland, USA) and was used as a control.

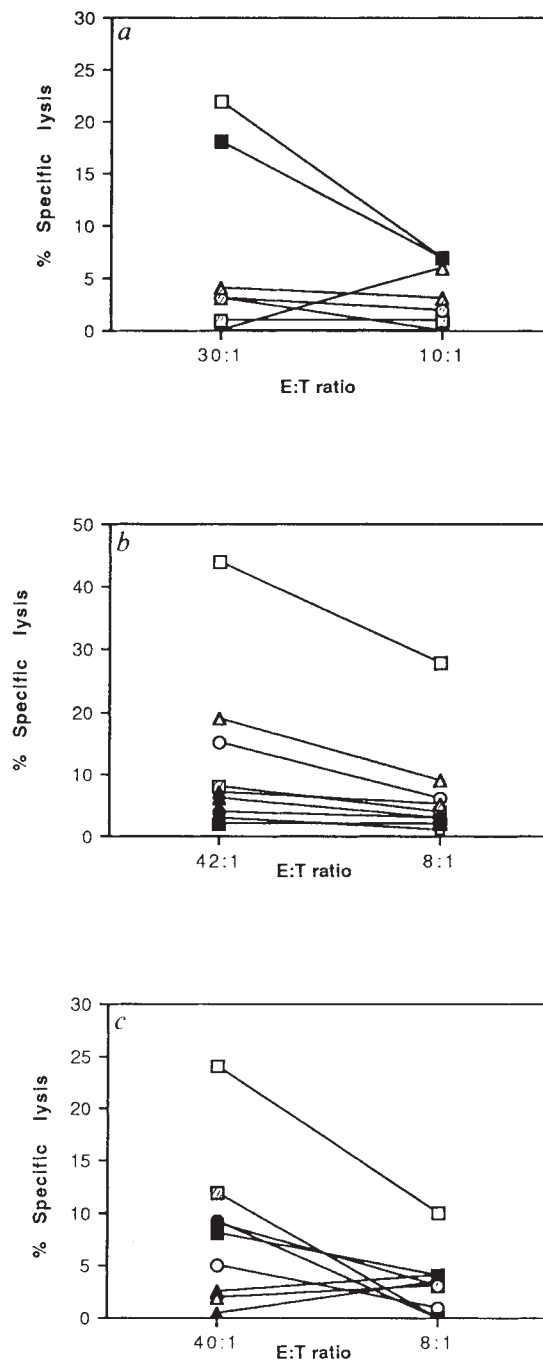


Fig. 2 HLA restriction of HIV gag-specific CTL. *a* and *b*, Effector CTL from donor 007 (HLA-A1, A32, B44, B27, DR1, DR4) were tested on HLA class I matched or mismatched target B lymphoblastoid cells. *a*, Targets were autologous (open symbols), shared HLA B27 only (closed symbols) or shared no HLA class I antigens (hatched symbols). *b*, Targets shared HLA-A1 only (open symbols), HLA B44 only (closed symbols) or no HLA class I antigens (hatched symbols). HLA types of targets: Donor 007, HLA-A1, A32, B44, B27, DR1, DR4; B27 match, HLA-A2, B27, DR2; A1 match HLA-A1, B8, B17, DR3, DR7; B44 match HLA-A2, B44, DRw6; class I mismatch, HLA-A2, B15, B51, DR4. *c*, Effector CTL from donor 008 (HLA-A2, B8, B14, DR2, DR13) were tested on HLA class I matched or mismatched target B lymphoblastoid cells. Targets shared HLA-A2 only (open symbols), HLA-B8 only (closed symbols) or no HLA class I antigen (hatched symbols). HLA types of targets: donor 008, HLA-A2, B8, B14, DR2, DR13; A2 match, HLA-A2, B15, B51, DR4; B8 match, HLA A1, B8, B17, DR3, DR7; class I mismatch HLA-A3, B7, DR2. Targets were infected with gag-vac (□, ■, ▨), M-vac (△, ▲, ▴) or were uninfected (○, ●, ⊙).

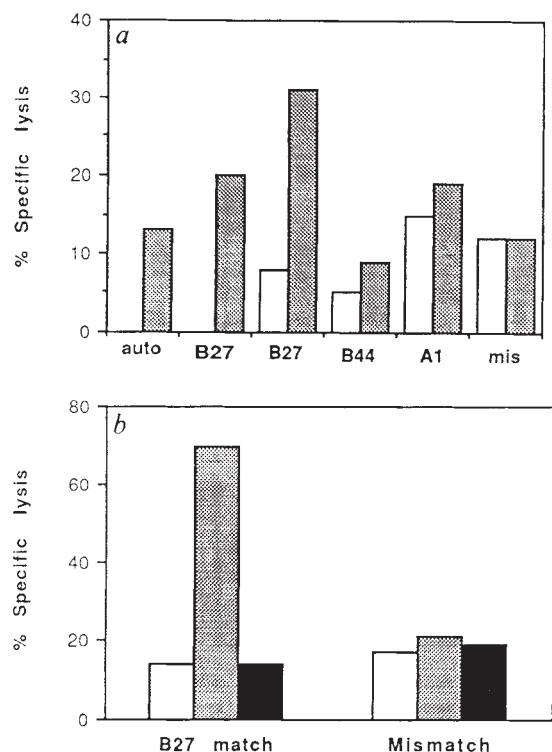
Table 1 Recognition of autologous and HLA class I mismatched targets by CTL from HIV seropositive donors

Donors	E:T	Targets						p24*	
		Autologous			HLA class 1 mismatch				
		gag	flu [†] (% specific lysis)	0	gag	flu [†] (% specific lysis)	0	Ab	Ag
001	32:1	49	22	12	16	10	6	NT	NT
002	22:1	12	0	0	2	0	1	+	NT
003	30:1	78	NT	0	8	0	1	NT	NT
004	10:1	9	NT	0	2	0	0	+	+
005	24:1	25	1	2	5	NT	10	+	—
006	22:1	29	NT	1	11	0	0	+	NT
007	8:1	45	19	15	2	4	4	+	NT
008	40:1	24	5	5	7	1	9	—	—
009	50:1	23	2	3	NT	NT	NT	NT	NT
010	24:1	50	NT	0	NT	NT	NT	—	—
011	72:1	47	19	15	NT	NT	NT	+	NT
012	40:1	18	5	0	NT	NT	NT	+	NT
013	72:1	21	7	0	NT	NT	NT	+	NT
014	30:1	6	4	3	2	1	0	+	NT
015	25:1	4	0	0	4	2	0	—	+

NT, not tested. Ab, antibody, Ag, antigen.

* p24 assays were performed using Abbott kits. Targets were Epstein-Barr virus transformed B-cell lines.

† Flu targets were infected with either M-vac or NP-vac which express influenza A virus matrix protein or nucleoprotein respectively.



The gag-specific CTL response of donor 007 was further examined by testing on autologous lymphoblastoid cells in the presence of a set of peptides that spanned the p24 sequence of gag. Peptides were added to the mixture of effector and target cells and only one was recognized, although the level of lysis was low (data not shown). We showed, using lymphoblastoid cells matched at single HLA alleles, that recognition of this peptide was restricted by HLA-B27 (Fig. 3a). As in previous experiments targets could be pre-incubated with peptide² and gave high levels of lysis (Fig. 3b). CTL lines could be grown in the presence of interleukin 2 on peptide-pulsed autologous irradiated lymphoblastoid feeder cells³ and lysed B27 matched target cells infected with gag-vac. By analogy with data accumulated for CTL recognition of influenza virus proteins²⁻⁴ we anticipate that other patients with HLA-B27 (which occurs in

Fig. 3 Effector CTL from donor 007 recognize a gag p24 peptide in association with HLA-B27. *a*, Effector CTL were tested on autologous, HLA class I matched or mismatched target B lymphoblastoid cells in the presence (hatched) or absence (open) of peptide 14 at a final concentration of 30 μ M. Targets were: autologous, shared HLA-B27 only (two different donors), HLA-B44 only, HLA-A1 only or HLA class I mismatched. E:T ratio 24:1. HLA types of targets: donor 007, HLA-A1, A32, B44, B27, DR1, DR4; B27 match HLA-A2, B27, DR2; B27 match, HLA-A2, B27, B60, DR11; B44 match, HLA-A2, B44, DRw6; A1 match, HLA-A1, B8, B17, DR3, DR7; class I mismatch HLA-A2, B15, B51, DR4. *b*, Effector CTL from 007 were tested on B27 matched or class I mismatched targets pre-incubated with no peptide (open), peptide 14 (hatched), or peptide 13 (closed). E:T ratio 6:1. HLA types of targets: B27 match, HLA-A2, B27, DR2; class I mismatch HLA-A2, B15, B51, DR4.

Methods. Peptides were made up in RPMI-1640-10% FCS. 10 μ l of each peptide solution was added to 150 μ l effectors/targets in a 96-well microtitre tray in duplicate. Specific lysis was measured in a 5-h ^{51}Cr release assay. Twenty-one overlapping 15- and 16-mer synthetic peptides corresponding to gag p24 sequence¹⁴ were manufactured by Cambridge Research Biochemicals under the MRC AIDS Directed Programme. Gag peptide 14 has the sequence:

265 K R W I I L G L N K I V R M Y C 279

The C-terminal cysteine was added to facilitate chemical coupling. Peptide preincubation method as described⁴. The peptide sensitized target cells in the concentration range 4×10^{-8} – 3×10^{-5} M.

7% of the population) should respond to the epitope and gag-specific CTL restricted by other HLA antigens should recognize different epitopes. It is likely that there are several other epitopes in the gag sequence and in other gene products. An epitope in the envelope glycoprotein that is recognized by H-2D^d murine CTL has been described⁹.

So far 15 HIV seropositive haemophiliac and homosexual patients have been tested for inducible gag-specific CTL responses (Table 1). Some of the responses were low but we consider that patients 001-013 gave positive responses. All responder patients were clinically well. Two of the patients had gag-specific CTL in the absence of p24 antibody, which has been shown to be a poor prognostic indication¹⁰. Long-term longitudinal studies of the spontaneous and inducible responses to gag and other viral products should help to define whether

CTL exert a beneficial or harmful role. This study shows that there is a gag-specific CTL response in nearly all healthy HIV seropositive patients. Other investigators have shown env- and pol-specific CTL responses in fresh PBM of some patients^{6,7}. Taken together, these results show that HIV-infected individuals make a potent CTL response to HIV antigens. CTL could control the persistent virus infection or destroy infected CD4-positive cells contributing to the immune dysfunction. If it can be shown that CTL exert a beneficial effect, efforts should be made to induce CTL memory responses by vaccination. It should be possible to find peptide epitopes restricted by the common HLA antigens using the methods described here; these could then be included in vaccines as peptides, fusion proteins or in other recombinant molecules.

This work was supported by a grant from the MRC AIDS Directed Programme. We are grateful to the staff of the Oxford Haemophilia Centre and of the Harrison Clinic for their help in collecting blood samples. We thank Professor P. Lachmann for making the gag p24 peptides available to us. Professor J. Edwards and David Hunter kindly provided some of the B lymphoblastoid cell lines used in the study. We thank Drs R. Eglin and J. Kurtz for advice and containment facilities, Joyce Macey for carrying out the p24 assays and Clare Crew for preparing the manuscript.

Received 28 June; accepted 1 November 1988.

1. Berman, P. W. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 5200-5204 (1988).
2. Townsend, A. R. M. *et al. Cell* **44**, 959-968 (1986).
3. McMichael, A. J., Gotch, F. M. & Rothbard, J. *J. exp. Med.* **164**, 1397-1406 (1987).
4. Gotch, F. M., Rothbard, J., Howland, K., Townsend, A. & McMichael, A. *Nature* **326**, 881-882 (1987).
5. Braciale, T. J. *et al. J. exp. Med.* **166**, 678-692 (1987).
6. Walker, B. D. *et al. Nature* **328**, 345-348 (1987).
7. Walker, B. D. *et al. Science* **240**, 64-66 (1988).
8. Plata, F. *et al. Nature* **328**, 348-351 (1987).
9. Takahashi, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 3105-3109 (1988).
10. Weber, J. *et al. Lancet* **i**, 119-122 (1987).
11. Hahn, B. H. *et al. Nature* **312**, 166-169 (1984).
12. Chakrabarti, S., Brechling, K. & Moss, B. *Molec. Cell. Biol.* **5**, 3403-3409 (1985).
13. Smith, G. L., Murphy, B. R. & Moss, B. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7155-7159 (1983).
14. Sanchez-Pescador, R. *et al. Science* **227**, 484-492 (1985).

Laminin receptor on platelets is the integrin VLA-6

Arnoud Sonnenberg*, Piet W. Modderman* & Frans Hogervorst†

* Central Laboratory of the Netherlands Red Cross Blood Transfusion Service and Laboratory for Experimental and Clinical Immunology, University of Amsterdam, The Netherlands
† The Netherlands Cancer Institute, Amsterdam, The Netherlands

Adhesion of platelets to the subendothelial matrix of an injured vessel wall is an essential step in triggering the formation of a haemostatic plug. Fibronectin, collagen and laminin are three major components of the subendothelial matrix which support platelet adhesion¹. Receptors for fibronectin and collagen have been identified on platelets and are included in the integrin family²⁻¹¹. Here we report that adhesion of platelets to laminin is inhibited by a rat monoclonal antibody against the integrin family member, VLA-6. This antibody does not affect platelet adhesion to fibrinogen, fibronectin or to type I and III collagen. Binding to laminin does not require platelet activation and is not inhibited by fibronectin and laminin cell-attachment peptides. Platelet adhesion to laminin is supported by Mn^{2+} , Co^{2+} and Mg^{2+} , but not by Ca^{2+} , Zn^{2+} and Cu^{2+} . This cation preference is distinct from that characteristic for other platelet-adhesive glycoproteins^{8,12,13}.

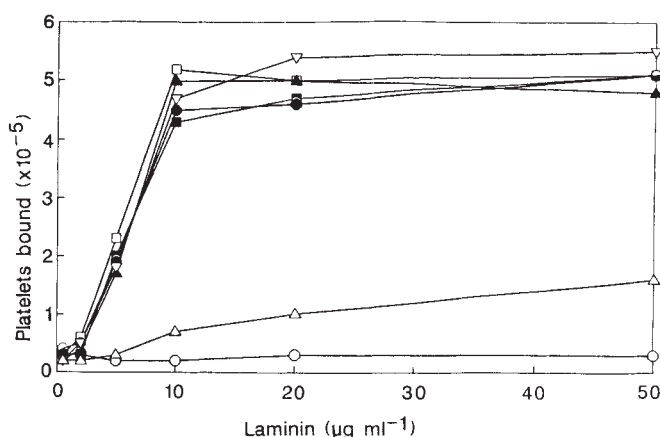


Fig. 1 Effect of monoclonal and polyclonal antibodies and of the synthetic peptide Gly-Arg-Gly-Asp-Ser on adhesion of platelets to microtitre plates coated with different concentrations of laminin. **Methods.** Microtitre plates (96-well, Greiner) were coated with laminin (Gibco-BRL) in phosphate-buffered saline (PBS) at 4°C for 12-18 h. Control wells were coated with bovine serum albumin (BSA). Each well was then washed twice with 0.5% BSA in PBS and incubated for 90 min in 1% BSA in PBS. Platelets were isolated from acid-citrate-dextrose anti-coagulated human blood and washed with a buffer³¹ containing 2 $\mu g\ ml^{-1}$ apyrase and 100 nM prostaglandin E_1 (PGE_1) at pH 6.5 to prevent platelet activation. Antibody-blocking experiments were performed essentially as described⁶. ^{51}Cr -labelled platelets at 1×10^8 per ml were preincubated in Iscove's medium containing 5 $mg\ ml^{-1}$ BSA and 40 nM PGE_1 for 60 min at room temperature with a 100-fold dilution of ascites fluid or serum, or with 1 mM Gly-Arg-Gly-Asp-Ser peptide (Bachem, Budendorf, Switzerland). For some experiments, the laminin-coated wells were preincubated with antibody and washed. Platelets were added to the wells and allowed to bind in the continued presence of antibody or peptide for 45 min at 37°C. The bound radioactivity was determined after the wells were washed. The number of platelets adhering to each well was calculated from the amount of radioactivity present in each well and the specific activity of the labelled platelets (3-5 c.p.m. per 1,000 platelets). Binding of platelets preincubated ○, with MAb GoH3, directed against VLA-6; ●, with MAb JsE3, a rat monoclonal antibody of the same IgG subclass as GoH3, but not reactive with human platelets; ▽, with Gly-Arg-Gly-Asp-Ser peptide. Both platelets and wells preincubated △, with polyclonal rabbit anti-laminin antibody. Binding to laminin-coated wells preincubated with the same antibody; ▲, with normal rabbit serum; □, with polyclonal goat anti-fibronectin antibody, or ■, with normal goat serum. No binding was observed on BSA-coated wells (not shown).

The integrins comprise a family of membrane receptors that function in cell-cell and cell-matrix interactions^{14,15}. Platelets express at least four distinct members of the integrin family: GPIIb-IIIa (ref. 16), VLA-2 (ref. 17), VLA-5 (ref. 11) and VLA-6 (ref. 11). GPIIb-IIIa is an activation-dependent receptor for RGD (Arg-Gly-Asp)-containing glycoproteins such as fibrinogen, von Willebrand factor and fibronectin^{3,4}. VLA-2 is a receptor for various types of collagens^{9,10} and VLA-5 serves as a receptor for fibronectin^{4,18,19}. The function of VLA-6 on platelets has not yet been identified. The VLA-6 protein is defined by monoclonal antibody (MAb) GoH3, previously shown to identify a complex of glycoproteins Ic and IIa on human and mouse platelets²⁰. GoH3 is directed against the Ic subunit (the α -subunit of VLA-6) of the Ic-IIa complex.

The presence of a laminin receptor on platelets was indicated by the finding that laminin supports platelet adhesion when coated on a solid substrate¹. Distinct receptors for laminin have been described for different cell types. These include a laminin-binding protein²¹⁻²³ of relative molecular mass 67K and ECMRI (ref. 24) (VLA-3), but the presence of the 67K laminin-binding protein on platelets has not been reported and platelets do not contain VLA-3 (ref. 11).

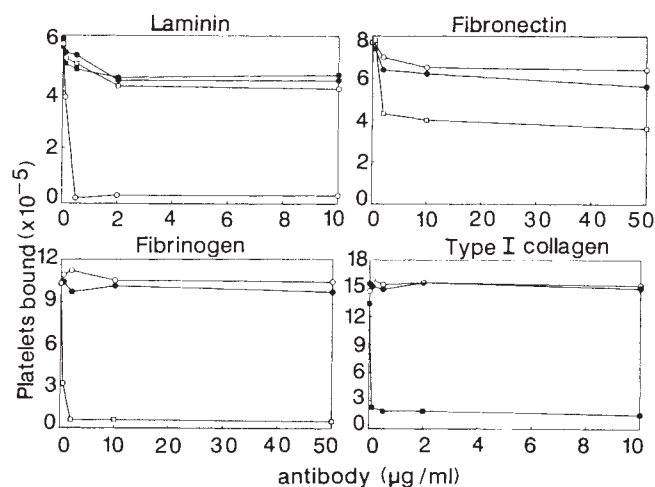


Fig. 2 Effect of various concentrations of monoclonal antibodies on adhesion of platelets to microtitre plates coated with laminin, fibronectin, fibrinogen or type I collagen. Microtitre plates were coated with 20 µg ml⁻¹ fibronectin (Sigma), laminin, fibrinogen (Kabi, Stockholm, Sweden) or type I collagen (Sigma). The effect of monoclonal antibodies (○, GoH3; ●, rat control antibody; □, C17; ■, 10G11) on platelet adhesion at the antibody concentrations indicated. All monoclonal antibodies were purified IgG fractions. The data are presented as the mean of duplicate determinations.

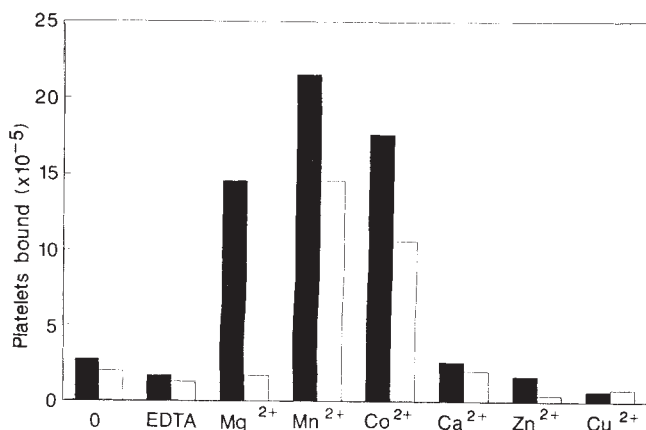


Fig. 3 Divalent cation-dependent adhesion of platelets to laminin and the effect of GoH3. Microtitre plates were coated with 25 µg ml⁻¹ laminin. The adhesion of ⁵¹Cr-labelled platelets to surface-coated microtitre wells was determined in the presence (□) or absence (■) of GoH3 antibody (2 µg ml⁻¹) in divalent cation-free medium (Tyrode's buffer containing 5 mg ml⁻¹ BSA and 40 nM PGE₁), in medium supplemented with 2 mM EDTA and in medium supplemented with 2 mM divalent cations, as indicated. The data are presented as the mean of quadruplicate determinations.

The possibility that VLA-6 is the laminin receptor on platelets was tested by observing the effect of the GoH3 antibody on platelet adhesion to laminin. As shown in Fig. 1, platelets adhered to surfaces coated with laminin in the absence of activation. Adhesion was dependent on the concentration of coating protein and was specifically blocked by polyclonal anti-laminin antibodies. Polyclonal anti-fibronectin and the fibronectin cell-attachment peptide Gly-Arg-Gly-Asp-Ser (refs 4 and 18) had no effect on adhesion. Neither did the synthetic peptide Cys-Asp-Pro-Gly-Tyr-Ile-Gly-Ser-Arg that blocks laminin-mediated cell adhesion to the 67K laminin-binding protein²⁵ inhibit platelet adhesion to laminin (not shown). The anti-VLA-6 antibody GoH3, however, prevented the adhesion of platelets to laminin (Fig. 1).

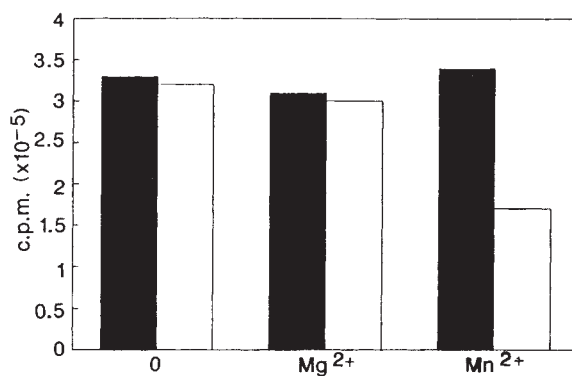


Fig. 4 The effect of divalent cations and soluble laminin on the binding of ¹²⁵I-labelled GoH3 antibody to platelets.

Methods. Platelets at 1×10^8 per ml were incubated at room temperature for 60 min with 1 µg ml⁻¹ ¹²⁵I-radiolabelled GoH3 antibody in the presence (□) or absence (■) of 100 µg ml⁻¹ soluble laminin in divalent cation-free medium (Tyrode's buffer containing 5 mg ml⁻¹ BSA and 40 nM PGE₁) or in medium supplemented with 2 mM divalent cations as indicated. Aliquots of the platelet suspensions (100 µl) were layered on to 700 µl 20% sucrose in Tyrode's buffer containing 0.5% BSA in a microcentrifuge tube and centrifuged for 3 min. Tips of the tubes were amputated and the amount of ¹²⁵I-labelled GoH3 associated with the platelet pellet was determined. The data are presented as the mean of triplicate determinations. Binding assays performed in the presence of 100 µg ml⁻¹ fibronectin gave results comparable to those obtained in the absence of added adhesive glycoproteins (not shown).

Additional antibody-blocking experiments (Fig. 2) were carried out to confirm that VLA-6 is a specific laminin receptor. Platelet adhesion to laminin, but not to fibronectin, fibrinogen or type I and III collagen (latter not shown), was blocked by MAb GoH3. Complete inhibition of adhesion to laminin was observed at 0.5 µg ml⁻¹ of GoH3. Adhesion to laminin was not affected by MAb C17 (anti-GPIIb-IIIa; ref. 26), 10G11 (anti-VLA-2; ref. 27) or a rat control antibody. The C17 and 10G11 antibodies, however, affected adhesion to other adhesive substrates. At 2 µg ml⁻¹ MAb C17 completely blocked adhesion to fibrinogen, partially inhibited adhesion to fibronectin and had little or no effect on adhesion to type I or type III collagen (Fig. 2). Adhesion of platelets to collagen types I and III was completely blocked by MAb 10G11 at 0.1 µg ml⁻¹. Adhesion to other adhesive substances was not affected by 10G11. These results are consistent with the identification of VLA-2 as a collagen receptor on platelets^{9,10}.

As the binding of all known integrins to their ligands is dependent on divalent cations, we tested their effect on VLA-6-mediated binding of platelets to laminin. Adhesion assays shown in Figs 1 and 2 were in Iscove's medium containing Ca²⁺ (1.5 mM) and Mg²⁺ (0.8 mM). As shown in Fig. 3, platelets adhered poorly to laminin in medium without divalent cations. Mn²⁺, Co²⁺ and Mg²⁺, but not Ca²⁺, Zn²⁺ and Cu²⁺, supported platelet adhesion at concentrations of 2 mM. This dependency is similar, but not identical, to that reported for platelet adhesion to collagen, which is supported by Mg²⁺, Mn²⁺, Co²⁺, Cu²⁺ and Zn²⁺ (ref. 8). Neither type of adhesion is supported by Ca²⁺. This might be a general feature of VLA-mediated adhesion and contrast with the cation requirements of GPIIb-IIIa (refs 12 and 13) and vitronectin receptor²⁸. Monoclonal antibody GoH3 at 2 µg ml⁻¹ effectively blocked Mg²⁺-dependent adhesion but only partially blocked Mn²⁺ and Co²⁺-dependent adhesion.

Binding of ¹²⁵I-labelled GoH3 to platelets was not itself affected by divalent cations. Thus, the inability of GoH3 to block platelet adhesion in the presence of Mn²⁺ and Co²⁺ is not due to decreased binding of GoH3. Soluble laminin

(100 $\mu\text{g ml}^{-1}$) blocked binding of ^{125}I -labelled GoH3 to platelets in the presence of Mn^{2+} (Fig. 4) and Co^{2+} (not shown), but not in the presence of Mg^{2+} . Therefore, we surmise that Mn^{2+} and Co^{2+} so increase the affinity of the receptor for laminin that GoH3 can no longer effectively compete for binding.

In summary, adhesion of platelets to laminin is mediated by VLA-6 and depends on divalent cations. Because it cannot be blocked by the Cys-Asp-Pro-Gly-Tyr-Ile-Gly-Ser-Arg peptide, the binding site for VLA-6 on laminin is probably distinct from that for the 67 K laminin-binding protein. A site present on the long arm of the laminin molecule is active in promoting neurite outgrowth^{29,30}. It will be of interest to see whether this site is the same as the binding site for VLA-6.

We thank Ms K. Linders and Ms R. Ganpat for their help with some of these experiments, Dr E. Ruoslahti for polyclonal anti-laminin and anti-fibronectin serum, and Drs H. Ploegh and A. E. G. Kr. von dem Borne for comments on the manuscript.

Received 25 July; accepted 11 October 1988.

1. Ill, C. R., Engvall, E. & Ruoslahti, E. *J. Cell Biol.* **99**, 2138–2145 (1984).
2. Gardner, J. M. & Hynes, R. O. *Cell* **42**, 439–448 (1985).
3. Plow, E. F., McEver, R. P., Collier, B. S., Woods, V. L. Jr & Marguerie, G. A. *Blood* **66**, 724–727 (1985).
4. Pytela, R., Pierschbacher, M. D., Ginsberg, M. H., Plow, E. F. & Ruoslahti, E. *Science* **231**, 1559–1562 (1986).

5. Giancotti, F. P. et al. *Blood* **69**, 1535–1538 (1987).
6. Piotrowicz, R. S., Orzechowski, R. P., Nugent, D. J., Yamada, K. Y. & Kunicki, T. J. *J. Cell Biol.* **106**, 1359–1364 (1988).
7. Nieuwenhuis, H. K., Akkerman, J. W. N., Houdijk, W. P. M. & Sixma, J. J. *Nature* **318**, 470–472 (1985).
8. Santoro, S. A. *Cell* **46**, 913–920 (1986).
9. Santoro, S. A., Rajpara, S. M., Staatz, W. D. & Woods, V. J. Jr *Biochem. biophys. Res. Commun.* **153**, 217–223 (1988).
10. Kunicki, T. J. et al. *J. Biol. Chem.* **263**, 4516–4519 (1988).
11. Hemler, M. E., Crouse, C., Takada, Y. & Sonnenberg, A. *J. Biol. Chem.* **263**, 7660–7665 (1988).
12. Bennett, J. S. & Viala, G. *J. Clin. Invest.* **64**, 1393–1401 (1979).
13. Marguerie, G. A., Edgington, T. S. & Plow, E. F. *J. Biol. Chem.* **255**, 154–161 (1980).
14. Hynes, R. O. *Cell* **48**, 549–554 (1987).
15. Ruoslahti, E. & Pierschbacher, M. D. *Science* **238**, 491–497 (1987).
16. Jennings, L. & Phillips, D. R. *J. Biol. Chem.* **257**, 10458–10463 (1982).
17. Pischel, K. D., Bluestein, H. G. & Woods, V. L. Jr *J. Clin. Invest.* **81**, 505–513 (1987).
18. Pytela, R., Pierschbacher, M. D. & Ruoslahti, E. *Cell* **40**, 191–198 (1985).
19. Takada, Y., Huang, C. & Hemler, M. E. *Nature* **326**, 607–609 (1987).
20. Sonnenberg, A., Janssen, H., Hogervorst, F., Calafat, J. & Hilgers, J. *J. Biol. Chem.* **262**, 10376–10383 (1987).
21. Rao, N. C., Barsky, S. H., Terranova, V. P. & Liotta, L. A. *Biochem. biophys. Res. Commun.* **111**, 804–808 (1983).
22. Malinoff, H. L. & Wicha, M. S. *J. Cell Biol.* **96**, 1475–1479 (1983).
23. Lesot, H., Kühl, U. & von der Mark, K. *EMBO J.* **2**, 861–865 (1983).
24. Wayner, E. A. & Carter, W. G. *J. Cell. Biol.* **105**, 1873–1884 (1987).
25. Graf, J. et al. *Cell* **48**, 989–996 (1987).
26. Modderman, P. W., Huisman, H. G., van Mourik, J. A. & von dem Borne, A. E. G. Kr. *Thromb. Haemostas* **60**, 68–74 (1988).
27. Giltay, J. C., Brinkman, H. J. M., Modderman, P. W., von dem Borne, A. E. G. Kr. & van Mourik, J. A. *Blood* (submitted).
28. Cheres, D. A. et al. *J. Cell Biol.* **105**, 1163–1173 (1987).
29. Edgar, D., Timpl, R. & Thoenen, H. *EMBO J.* **3**, 1463–1468 (1984).
30. Engvall, E. et al. *J. Cell Biol.* **103**, 2457–2465 (1986).
31. Legrand, C., Dubernard, V. & Nurdén, A. T. *Biochim. biophys. Acta* **812**, 802–810 (1985).

Abdominal segmentation of the *Drosophila* embryo requires a hormone receptor-like protein encoded by the gap gene *knirps*

Ulrich Nauber*†, Michael J. Pankratz*†, Andrea Kienlin†, Eveline Seifert†, Uwe Klemm‡ & Herbert Jäckle*

* Institut für Genetik und Mikrobiologie, LMU München, Maria Ward Strasse 1a, D-8000 München 19, FRG

† Max Planck Institut für Entwicklungsbiologie, Spemannstrasse 35, D-7400 Tübingen, FRG

The body pattern along the anterior–posterior axis of the insect embryo is thought to be established by two organizing centres localized at the ends of the egg¹. Genetic analysis of the polarity-organizing centres in *Drosophila* has identified three distinct classes of maternal effect genes that organize the anterior, posterior and terminal pattern elements of the embryo². The factors provided by these gene classes specify the patterns of expression of the segmentation genes at defined positions along the longitudinal axis of the embryo^{3,4}. The system responsible for organizing the posterior segment pattern is a group of at least seven maternal genes² and the zygotic gap gene *knirps* (*kni*). Their mutant phenotype has adjacent segments in the abdominal region of the embryo deleted². Genetic analysis and cytoplasmic transplantation experiments suggested that these maternal genes are required to generate a 'posterior activity' that is thought to activate the expression of *kni* (reviewed in ref. 2). The molecular nature of the members of the posterior group is still unknown. Here we report the molecular characterization of the *kni* gene that codes for a member of the steroid/thyroid receptor superfamily of proteins which in vertebrates act as ligand-dependent DNA-binding transcription regulators.

The phenotypes produced by several *kni* alleles can be ordered into a phenotypic series (Fig. 1). The head, the thorax, the eighth abdominal segment and the tail region of homozygous *kni*

embryos are normal, but different numbers of adjacent abdominal segments are deleted in the various *kni* mutant alleles (for a description of the phenotypic series see Fig. 1 and ref. 5). Chromosomal rearrangements that are associated with a *kni* mutation enabled the *kni* locus to be localized to region 77E in the left arm of the third chromosome (Fig. 2a). This chromosome region was microcloned and used to initiate a chromosomal walk between the chromosome bands 77E and 77F (Fig. 2b, c). To identify the *kni* locus on a molecular scale, chromosomal breakpoints defining the *kni* region were mapped (Fig. 2a). Sequences encoding essential *kni* gene functions were identified by analysis of the X-ray-induced allele *kni*^{FC13} which produces an amorphic *kni* phenotype. This *kni* allele is associated with a small deficiency of about 2 kilobase (kb) of DNA (Fig. 2b) that encodes a portion of an embryo-specific transcription unit. Northern blot hybridization (Fig. 2d) with a corresponding DNA probe showed two zygotic transcripts of 2.2 and 2.5 kb, which code for the same protein (see below). These transcripts are not present in *kni*^{FC13} homozygous embryos (data not shown). The smaller transcript is expressed only transiently at the blastoderm stage, although the larger transcript continues to be expressed after gastrulation. The developmental profile of expression of the 2.2 kb transcript is therefore consistent with the expected expression period of *kni* for segmentation function⁵.

To demonstrate that *knirps* protein function is encoded by this embryo-specific transcription unit, we injected the putative *kni* antisense RNA into wild-type embryos to phenocopy the *kni* mutant phenotype. As shown in Table 1, 5% of the embryos developed *kni* phenocopies. The segmentation defects of these embryos were similar to those of weak *kni* alleles (Fig. 1); the strongest phenocopy had a deletion of four abdominal segments (Fig. 1f). However, defects comparable to extreme *kni* phenotypes were never observed. The production of *kni* phenocopies was independent of the site of injection (Table 1), and control injections with sense RNA had no significant effect on the segmentation pattern of the embryos. These experiments, in combination with the detection of a small *kni* deficiency which maps into the transcribed region (see above) indicated that these transcripts encode the *kni* gene product.

Sequence analysis of a 2 kb *kni* complementary DNA clone and of its genomic counterpart revealed three exons which are interrupted by two introns of 733 base pairs (bp) and 214 bp.

† Present address: Institut für Humangenetik, Gosslerstrasse 12d, D-3400 Göttingen, FRG.



Fig. 1 The *kni* phenotypic series and a *kni* phenocopy produced by injection of *kni* antisense RNA. *a*, Wild-type *Drosophila* larva. A1-8, abdominal segments 1-8. *b-d*, Intermediate and weak *kni* phenotypes¹⁶ produced by *kni*^{14F} *b*, *kni*^{5F} (*c*) and *kni*¹ alleles with only three to five abdominal segments. *e*, Extreme *kni* phenotype (homozygous *kni*^{FC13} embryo¹⁶). Segments A1-7 are fused and replaced by a single segment which shows a broad denticle field on the ventral side. Segment A8 is normal. *f*, *kni* phenocopy produced in a wild-type embryo after injection of *kni* antisense RNA; this phenotype is similar to that of a weak *kni* allele.

Methods. Preparation of cuticles was as described^{17,18}. For details of *kni* antisense RNA injections see Table 1.

Table 1 Phenocopying of *kni* resulting from injection of *kni* antisense and sense RNA into *Drosophila* wild-type embryos

Injection of antisense RNA					
Site of injection	Number of injected embryos	Developed embryos	Hatched larvae (%)	Injection artefacts (%)	<i>kni</i> phenocopies (%)
Anterior	70	62	48 (77)	9 (15)	5 (8)
50% EL	120	113	66 (58)	44 (39)	3 (3)
Posterior	120	114	83 (73)	25 (22)	6 (5)
Total	310	289	197 (68)	78 (29)	14 (5)
Injection of sense RNA					
Anterior	40	35	31 (89)	4 (11)	0
50% EL	120	100	81 (81)	19 (19)	0
Posterior	120	102	82 (80)	20 (20)	0
Total	280	237	194 (82)	43 (18)	0

For the production of *kni* sense and antisense RNA a 2-kb cDNA clone (cJ15; see Fig. 3) was subcloned into the *Eco*RI site of 'Bluescript' (Stratagene) and RNA was transcribed *in vitro* from each strand of the cDNA using the T3 and T7 promoter sequences of the 'bluescript' vector and T3 and T7 RNA polymerase, respectively. The preparation of *kni* sense and antisense RNA, determination of the RNA concentration and alkaline hydrolysis to a mean length of about 200 bases was essentially as described¹⁵, with minor modifications in the use of T3 and T7 RNA polymerase (according to the manufacturer's protocol for the 'bluescript' vector). Wild-type embryos were injected as described¹⁵ before pole cell formation at different sites with about 300 pl of *kni* sense or antisense RNA (concentration 8 mg ml⁻¹). % EL refers to egg length; posterior pole is 0%, anterior tip is 100%.

A single open reading frame of 1,287 bp would encode a basic protein (estimated *pI* = 10.2) of 429 amino acids with a calculated relative molecular mass of 45,600 (45.6K) (Fig. 3). Sequence comparison with other known proteins revealed a striking similarity of the N terminus of the knirps protein with the DNA-binding domain of members of the steroid/thyroid hormone receptor superfamily (Fig. 4) which include the receptors for vitamin D and the vitamin A-related morphogen retinoic acid (reviewed in ref. 6). Between 45 and 53% of the amino acids of this domain (67 residues in *kni*) are conserved among the different members of the receptor family. In all of them, including *kni*, 20 amino acids are invariant within the conserved region. This region is characterized by precisely spaced Cys residues which could potentially fold into a 'finger-like' structure (reviewed in ref. 6). This is reminiscent of the 'zinc fingers' of the other gap genes *hunchback* (*hb*)⁴ and (*Krüppel*) *Kr*⁷ and several known transcription activators (for a review see ref. 8).

It is therefore likely that *kni*, like *hb* and *Kr*, encodes a potential DNA-binding protein. In contrast to them, however, *kni* encodes the DNA-binding motif characteristic of the superfamily of ligand-dependent nuclear receptor molecules⁶. This superfamily of DNA-binding transcription activators includes hormone receptors and non-hormone responsive members such as the receptor for retinoic acid (for a review see ref. 6) which acts as a morphogen by generating the anterior-posterior pattern in the developing chick limb bud⁹ and by inducing parietal endoderm in mouse teratocarcinoma cell lines¹⁰. The involvement of the retinoic acid receptor molecule in the pattern formation processes of vertebrates and the fact that the Cys/Cys finger motif has so far only been found in ligand-responsive nuclear receptor molecules may suggest that the *kni* activity depends on a ligand molecule as well. This ligand could be the active component generated by the maternal group of genes that organize the posterior segment pattern. Information encoded by these genes is initially localized at the posterior pole, and it moves

Fig. 4 Similarity between the predicted *kni* protein and DNA-binding domains of several members of the steroid/thyroid nuclear receptor superfamily. Comparison with the amino-acid sequence of the human receptors for oestrogen (hER)²², thyroid hormone (hTR α ; hTR β)²⁵⁻²⁷, v-erbA²⁸, retinoic acid (hRAR α ; hRAR β)^{22,29-32}, glucocorticoid (hGR)^{33,34}, progesterone (hPR)³⁵, mineralocorticoid (hMR)³⁶, androgen (hAR)^{37,38}, and vitamin D (hVDR)³⁹. Numbers on the left correspond to the numbering of amino acids, percentages on the right side indicate the extent of homology with the *kni* protein. The conserved region consists of 66-68 amino acids. The region compared is indicated by lines on top and bottom. Conserved amino acids are marked by asterisks. *** Consensus sequence of invariant amino acids in all of the proteins shown.

<i>kni</i>	1	MQQCKVCGEPAAGFHGAFTCEGCKSFFGRSYN NIS TISECKNEGKCIIDKKNRITTCACRLKRCYNVGMSSGGSRYGRR	
hER	181	ETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQ GHN DYM CPATNQCTIDKNRRKSCQACRLKCYFVGMHKGIRKDRR	50 %
		* * * * *	
hTR α	49	KDEQCVVCGDKATGYHYRCITCEGCKGFFRTIQNLHPTYS CKYDSCCVIDKITRNQCQLCFKKCIJAVAMAMDLVLDDSK	50 %
		* * * * *	
hTR β	98	KDELVCVCGDKATGYHYRCITCEGCKGFFRTIQNLHPTYS CKYEGKCVIDKITRNQCQCFKKCIYVGMATDLVLDDSK	53 %
		* * * * *	
v-erbA	33	KDEQCVVCGDKATGYHYRCITCEGCKSFFRTIQNLHPTYS CTYDGCVCVIDKITRNQCQLCFKKCIJAVAMAMDLVLDDSK	51 %
		* * * * *	
hRAR α	54	IYKPCFVCGDKSSGYHYGVSAECGCKGFFRSIQNM VYT CHRDKNCILNKVTRNRCQYCRKQCFEVGMSSKESVRNDRD	47 %
		* * * * *	
hRAR β	77	VYKPCFVCGDKSSGYHYGVSAECGCKGFFRSIQNM IYT CHRDKNCVINKVTRNRCQYCRKQCFEVGMSSKESVRNDRN	45 %
		* * * * *	
hGR	417	PPKLCIVCSDEASGCHYGVLTCGCKSVFFKRAVE GQH NYL CAGRNDICIDKIRKNCPCACRYKCLQAGMNLKARKTKKK	45 %
		* * * * *	
hPR	563	PQRICLICGDEASGCHYGVLTCGCKSVFFKRAVE GQH NYL CAGRNDICIDKIRKNCPCACRYKCLQAGMNLGGRKPKKP	45 %
		* * * * *	
hMR	599	PSKICLVCGDEASGCHYGVLTCGCKSVFFKRAVE GQH NYL CAGRNDICIDKIRKNCPCACRYKCLQAGMNLGARKSKKL	47 %
		* * * * *	
hAR		KTCLICGDEASGCHYGALTGCKSVFFKRAAE GKQ KYL CASRNDCTIDKFRKNCPCACRYKCYEAGMTLGA	47 %
		* * * * *	
hVDR	20	VPRICVCGDRATGFGHFNAMTCEGCKGFFRSIMK RKA LFT CFPNGDCRITKDNRRHCQACRLKRCVDIGMKEFILTDEE	53 %
		* * * * *	
***		C-C---G-H---C-CK-FF-R-----C---C---K-R-R-C-CR---C---GM	

anteriorly in the developing embryo before *kni* is first expressed. The details of the localization and translocation process as well as the molecular form of the translocated material are at present unclear. But it is known from transplantation experiments (reviewed in ref. 2) that the material is required in the presumptive abdominal region of the embryo, and that one member of the posterior group of genes, *pumilio*, seems to be involved in the translocation of the material, possibly the *nanos* gene product, from the pole to its site of action¹¹. The gene *nanos*, likely to be the active component of the posterior organizer centre, could therefore be the ligand that activates the *kni* protein in the abdominal region. This hypothesis was tested by *in situ* hybridization of the *kni* probe to tissue sections of wild-type and maternal effect mutant embryos. But the lack of *kni* expression in the posterior part of embryos laid by homozygous *nanos* females (U.N. *et al.*, manuscript in preparation) clearly indicates that localized expression of *kni* in the presumptive abdominal region is dependent on the presence of the *nanos* gene product. Therefore, the control of *kni* activity by the posterior group of genes seems to occur at the transcription level, as for *hb* and *Kr*^{12,13}, rather than by activating the *kni* gene product.

Except for the conservation of the Cys/Cys finger motif of the hormone receptor superfamily of proteins, which puts the *kni* protein into the class of DNA-binding proteins, there is no

further evidence to support ligand-dependent activation of the *kni* protein. Thus, *knirps* may represent a first exception among the Cys/Cys finger proteins which acts in a ligand-independent manner. In this view, *kni* may have functions analogous to *hb* and *Kr*, the other gap genes, in the abdominal region of the embryo. The information inherent in their domains of activity is used in two distinct processes. By repression and/or activation of subordinate genes, it eventually regionalizes the embryo by delimiting the domains of homeotic gene expression. In addition, this information is used to regulate members of the pair-rule class of genes, whose periodic patterns of expression precede the metamorphosis of the embryo (reviewed in ref. 14). It is not clear why the *kni* protein is slightly, but maybe importantly, different from the *hb* and *Kr* proteins, and possibly also from the other members of the superfamily of hormone receptors.

We would like to thank Dr R. Lehmann for providing the basis for this study, Drs S. Cohen, U. Gaul, D. Tautz and G. Jürgens for their important contributions. The work was supported by the Genzentrum München, a grant of the Deutsche Forschungsgemeinschaft (Leibniz Programm) and by the Max-Planck-Gesellschaft. The sequence data in this publication have been submitted to the EMBL/GenBank Libraries. These sequence data will appear in the EMBL/GenBank/DBJ Sequence Database under accession number X13331.

Received 29 July; accepted 5 October 1988.

- Sander, K. *Adv. Insect Physiol.* **12**, 125-238 (1976).
- Nüsslein-Volhard, C., Frohnhöfer, H.-G. & Lehmann, R. *Science* **238**, 1675-1681 (1987).
- Knipple, D. C. *et al. Nature* **317**, 40-44 (1985).
- Tautz, D. *et al. Nature* **332**, 281-284 (1987).
- Lehmann, R. *Development* (in the press).
- Evans, R. M. *Science* **240**, 889-895 (1988).
- Rosenberg, U. B. *et al. Nature* **319**, 336-339 (1986).
- Evans, R. M. & Hollenberg, S. N. *Cell* **52**, 1-3 (1988).
- Thaller, C. & Eichele, G. *Nature* **227**, 625-627 (1987).
- Strickland, S. & Mahdavi, V. *Cell* **15**, 393-395 (1978).
- Lehmann, R. & Nüsslein-Volhard, C. *Nature* **329**, 167-170 (1987).
- Tautz, D. *Nature* **332**, 281-284 (1988).
- Gaul, U. & Jäckle, H. *Cell* **51**, 549-555 (1987).
- Akam, M. E. *Development* **101**, 1-22 (1987).
- Rosenberg, U. B., Preiss, A., Seifert, E., Jäckle, H. & Knipple, D. C. *Nature* **313**, 703-706 (1985).
- Lehmann, R. thesis, Univ. Tübingen (1985).
- Wieschaus, E., Nüsslein-Volhard, C. & Kluding, H. *Dev. Biol.* **104**, 172-186 (1984).
- Preiss, A., Rosenberg, U. B., Kienlin, A., Seifert, E. & Jäckle, H. *Nature* **313**, 27-32 (1985).
- Pirrotta, V., Jäckle, H. & Edström, J. E. in *Genetic Engineering: Principles and Methods* Vol. 5, 1-17 (Plenum, New York, 1983).

- Hafen, E., Basler, K., Edström, J. E. & Rubin, G. M. *Science* **236**, 55-63 (1987).
- Mount, S. M. *Nucleic Acids Res.* **10**, 459-472 (1982).
- Petkovich, M., Brand, N. J., Krust, A. & Chambon, P. *Nature* **330**, 444-450 (1987).
- Zahraoui, A. & Cuny, G. *Eur. J. Biochem.* **166**, 63-69 (1987).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Benbrook, D. & Pfahl, M. *Science* **238**, 788-791 (1987).
- Nakai, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 2781-2785 (1988).
- Weinberger, C. *et al. Nature* **324**, 641-646 (1986).
- Debuire, B. *et al. Science* **224**, 1456-1459 (1984).
- Giguere, V., Ong, E. S., Segui, P. & Evans, R. M. *Nature* **330**, 624-629 (1987).
- de The, H., Marchio, A., Tiollais, P. & Dejan, A. *Nature* **330**, 667-670 (1987).
- Brand, N. *et al. Nature* **332**, 850-853 (1988).
- Benbrook, D., Lernhardt, E. & Pfahl, M. *Nature* **333**, 669-672 (1988).
- Hollenberg, S. M. *et al. Nature* **318**, 635-641 (1985).
- Govindan, M. V., Devic, M., Green, S., Gronemeyer, H. & Chambon, P. *Nucleic Acids Res.* **13**, 8293-8304 (1985).
- Misrahi, M. *et al. Biochem. biophys. Res. Commun.* **143**, 740-748 (1987).
- Ariza, J. L. *et al. Science* **237**, 268-275 (1987).
- Chang, C., Kokontis, J. & Shutsung, L. *Science* **240**, 324-326 (1988).
- Lubahn, D. B. *et al. Science* **240**, 327-330 (1988).
- Baker, A. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 3294-3298 (1988).

The *Drosophila* gene *knirps-related* is a member of the steroid-receptor gene superfamily

Anthony E. Oro^{*†}, Estelita S. Ong^{*},
Jonathan S. Margolis[†], James W. Posakony[†],
Michael McKeown[‡] & Ronald M. Evans^{*}

^{*} Howard Hughes Medical Institute, Gene Expression Laboratory, Salk Institute for Biological Studies, La Jolla, California 92038, USA
[†] University of California at San Diego, Department of Biology B-022, and Center for Molecular Genetics, La Jolla, California 92093, USA

[‡] Molecular Biology and Virology Laboratory, Salk Institute for Biological Studies, La Jolla, California 92038, USA

Molecular cloning has demonstrated that the receptors for steroid, retinoid and thyroid hormones are part of a large superfamily of nuclear regulatory proteins. In vertebrates these molecules regulate diverse biological processes such as pattern formation, cellular differentiation and homeostasis¹. The universal necessity for embryonic and adult cells to respond to their external environment might mean that members of this family pre-date the divergence of vertebrates and invertebrates. We have screened a *Drosophila* genomic library for steroid receptor homologues using a human retinoic acid receptor complementary DNA^{2,3} as a hybridization probe. Several clones were recovered, one of which mapped to chromosomal position 77E1-2, the cytological location of the gap segmentation gene *knirps*⁴. Sequence analysis of a cDNA clone representing the human retinoic acid receptor homologue showed similarity of the predicted protein to the vertebrate steroid receptors, as well as to the predicted *knirps*⁵ gene product. *In situ* hybridization of a cDNA probe to wild-type embryos revealed a uniform distribution of transcripts that were apparently maternally derived. Zygotic transcript accumulation begins in a broad antero-ventral domain before the cellular blastoderm stage. At the cellular blastoderm stage two additional circumferential bands of transcript appear.

To identify potential homologues of the vertebrate steroid receptors, a Southern blot of *Drosophila* genomic DNA was probed with a cDNA fragment encoding the human retinoic acid receptor (hRAR) DNA-binding domain^{2,3}. Under conditions of reduced hybridization stringency, six distinct *Eco*RI bands ranging in size from 2 kilobases (kb) to larger than 12 kb were detected (Fig. 1a, lane 1). Screening a *Drosophila* genomic library with the same probe and using the same hybridization conditions resulted in the isolation of three classes of inserts (based on cross-hybridization under high-stringency condi-

tions). Representatives of each class were hybridized to larval salivary gland polytene chromosomes to identify their cytogenetic location. One class of inserts maps to 77E1-2, the same location as that identified for the gap segmentation gene *knirps* (*kni*)⁴. A portion of the genomic insert that hybridized most strongly to the hRAR probe was subcloned and sequenced (plasmid pRX4.5). The amino-acid sequence derived for one of the reading frames contains the structural features of a steroid-receptor DNA-binding domain (Fig. 1b). Moreover, this sequence has a striking similarity to the predicted amino-acid sequence of the *kni* product⁵. This new locus will be referred to as *knirps-related* (*knrl*).

To characterize transcripts from the *knrl* locus, the genomic fragment pRX4.5 was used as a probe to screen a total third instar larval imaginal disc cDNA library. Three cDNA clones were isolated and the complete sequence of one insert, the 3,505-base pair (bp) Imd2, is shown in Fig. 2. Imd2 contains an open reading frame capable of encoding 647 amino acids, beginning with the presumptive initiator methionine at nucleotide 517, and ending with a stop codon beginning at position 2,458.

A comparison of the predicted *knrl*-encoded protein with other members of the steroid/thyroid receptor superfamily is shown in Fig. 3. First, sequence alignment demonstrates that the greatest similarity with the other receptors lies in the 67 amino acids of the putative *knrl* DNA-binding domain (Fig. 3a). Between amino acids 14 and 80 of *knrl* there is 85% amino acid identity with the *kni* product, 49% with the human thyroid receptor, 47% with the human retinoic acid receptor and 43% with the human glucocorticoid receptor. Interestingly, the *knrl* and *kni* DNA-binding domains both contain a glycine in the region linking the two zinc fingers (residues 39 and 30 in *knrl* and *kni* respectively) at a position which in all other receptors¹ contains an arginine or a lysine. This implies that these two genes could have a common origin. Secondly, the *knrl* carboxy-terminal sequence shares little similarity with the same region of the other receptors. The carboxyl terminus of vertebrate receptors contains the ligand-binding function, and it has been shown that relatedness between their carboxyl termini roughly reflects the relatedness of their ligand structures^{1,6}. Therefore a putative *knrl* ligand would probably differ from the steroid, retinoid or thyroid hormone classes of ligands.

Analysis of the temporal and spatial expression of *knrl* suggests that it may function both in early embryogenesis and throughout later development. A Northern blot of stage-specific RNA showed that a single RNA species of ~3.8 kb which is expressed at low levels between 0–3 h after egg-laying increased significantly in later embryos, and also in larvae and adults (data not shown). The spatial location of *knrl* transcripts was assayed by *in situ* hybridization using *knrl* antisense RNA on sections

Fig. 1 Cross-hybridizing bands reveal the presence of retinoic acid receptor-related sequences in *D. melanogaster*. **a**, Low-stringency hybridization using the 800 bp *Eco*RI-*Sac*I fragment of an hRAR cDNA as a probe shows six hybridizing *Eco*RI bands (lane 1) and five *Xho*I bands (lane 2), but none at high stringency. A *Drosophila* genomic library was screened; one class of isolates contained the 4.5 kb *Eco*RI-*Xba*I genomic fragment subcloned in pRX4.5. pRX4.5 hybridizes to the 14-kb *Eco*RI band and the 11-kb *Xho*I band at high stringency, which is indicative of a unique gene. **b**, Map of subclone pRX4.5 showing the smallest hybridizing fragment, an 800-bp *Pst*I-*Dra*I fragment, as a hatched box. Nucleotide sequencing showed a region homologous with vertebrate steroid hormone receptor genes. The predicted amino-acid sequence is characteristic of the second zinc finger of a steroid hormone receptor. Restriction enzyme cutting sites are R, *Eco*RI; D, *Dra*I; K, *Kpn*I; S, *Sac*I; B, *Bam*HI; Bg, *Bgl*II; Xb, *Xba*I. Standard methodology was used⁹; conditions for hybridization at low and high stringency were as described elsewhere¹⁰.

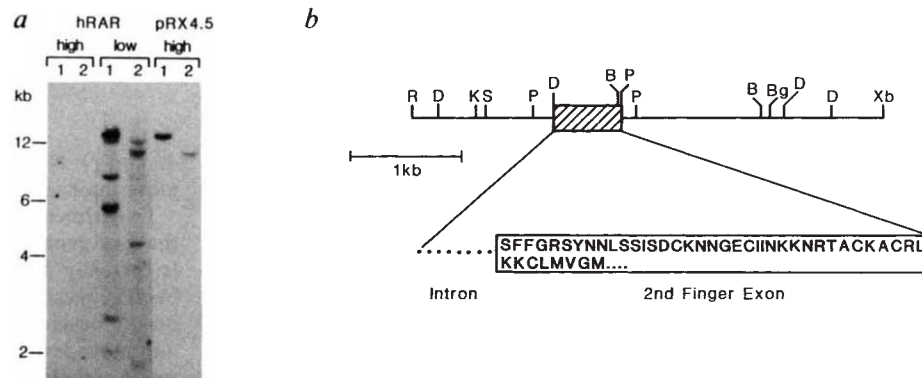


Fig. 3 Comparison of the predicted *knrl* product with vertebrate steroid/thyroid hormone receptors. *a*, Alignment of the DNA-binding domains of representative members of the superfamily, showing the conserved amino acids and the extensive structural similarity between the predicted *knrl* and *kni* proteins. Note that the identity of *knrl* and *kni* proteins extends past the conserved Gly and Met residues of the DNA-binding domain. *b*, Overall structural comparison of the predicted protein sequence of *knrl* with other members of the steroid/thyroid hormone receptor superfamily. The region marked 'DNA' is compared with the 66–68 amino acid DNA-binding domains, and the region marked 'ligand binding' is compared with the amino acids starting at a point 25 amino acids past the conserved Gly and Met residues of the DNA-binding domain. The intervening 25 amino acids past the conserved Gly and Met are 94% identical in *knrl* and *kni*, but are not significantly related to the other three vertebrate receptors (indicated by an asterisk). As there is no significant similarity between *knrl* and the other receptors in their carboxy-terminal regions, no specific alignment of these regions is shown. The programs of Devereux *et al.*¹² were used for comparisons. Numbers indicating amino acids are as detailed for *knrl* (this report), *kni*⁵, *hT₃R β* ¹³, *hRAR α* ² and *hGR*¹⁴.

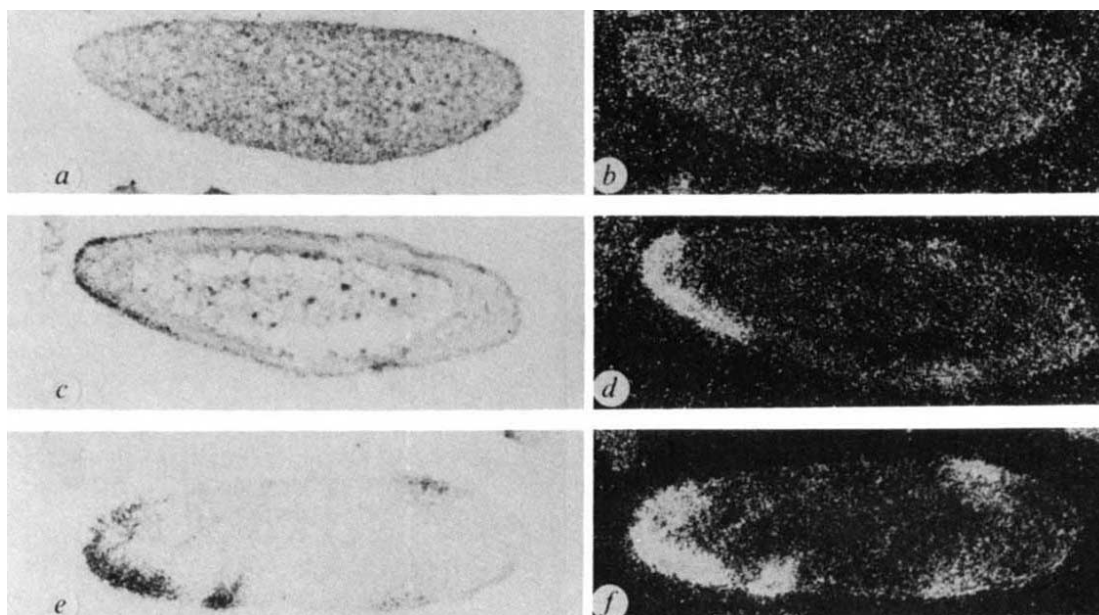
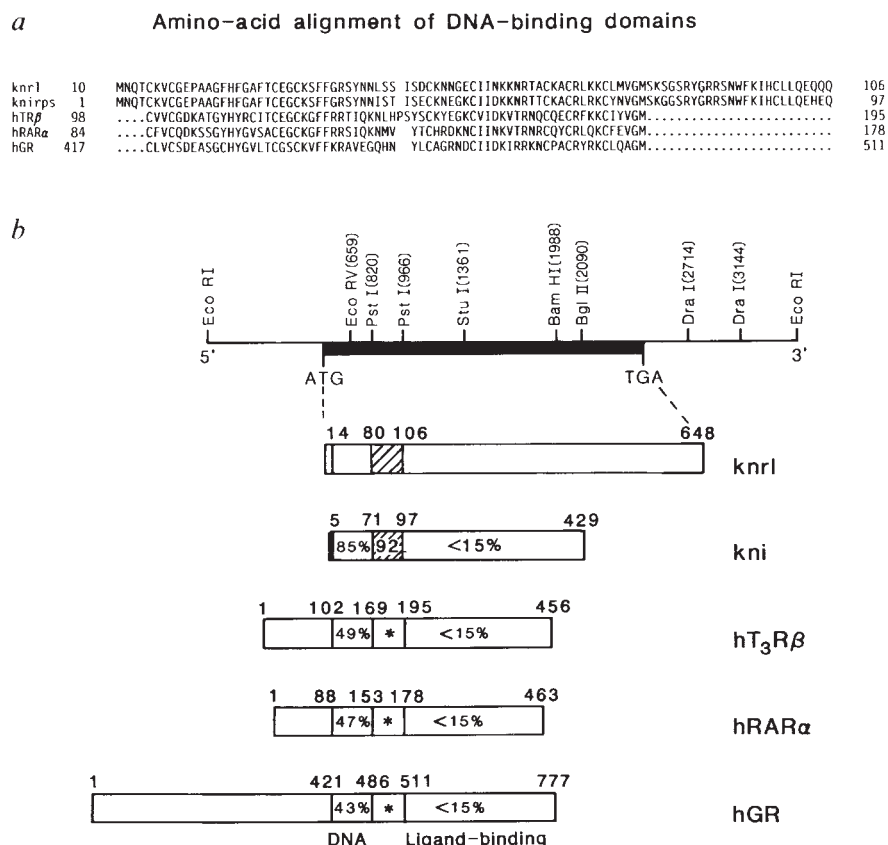


Fig. 4 Spatial localization of *knrl* transcripts in early embryos. Visualization of *knrl* transcripts by *in situ* hybridization to sections of wild-type *D. melanogaster* embryos shows both maternal and zygotic expression. As a spatial and temporal marker, alternate sections were hybridized with a probe representing the *Drosophila* gap gene *hunchback*¹⁵ (data not shown). Embryos are oriented with anterior to the left and dorsal at the top; staging as in ref. 16. *a*, *b*, Cleavage stage embryo showing the spatially uniform distribution of apparent maternal *knrl* transcripts. *c*, *d*, Embryo at syncytial blastoderm showing apparent early zygotic *knrl* expression in an antero-ventral domain extending from 80 to 100% of egg length along the ventral side of the embryo. Also apparent is the first appearance of the most posterior of the three cellular blastoderm expression domains (see below). *e*, *f*, Cellular blastoderm pattern of *knrl* expression. The slightly oblique section shows intense antero-ventral transcript accumulation (domain I), as well as the two more posterior transcript stripes at ~70% (domain II) and 25% (domain III) of egg length ventrally. Transcript accumulation in domain II is always significantly higher ventrally than it is dorsally. *In situ* hybridization was essentially as described¹⁷ with modifications¹⁸. [³⁵S]UTP was incorporated into antisense RNA from the T7 promoter of a Stratagene Bluescript plasmid containing the insert of the cDNA clone *Imd2*.

hitherto unrecognized small-molecule morphogen, and some of the genes involved in regulating *knrl* function might affect the synthesis of the ligand or storage of a ligand precursor, rather than regulating *knrl* expression. However, the dissimilarity of the carboxyl termini of the *knrl* gene product and of the other receptors makes it difficult to predict a potential ligand. Perhaps the *knrl*⁺ product is a constitutive transcriptional regulator and functions without a ligand. But in any case, *knrl* belongs to the steroid receptor gene family whose members contribute to morphogenesis and pattern formation in both invertebrates and vertebrates.

The authors thank Herbert Jackle for sequence information before its publication; Ruth Lehmann, Bill Segraves, David Mangelsdorf and Eva Lee for discussion, Charles Zuker for reagents and Elaine Stevens for administrative assistance.

Received 21 October; accepted 4 November 1988.

1. Evans, R. M. *Science* **240**, 889-895 (1988).
2. Giguere, V., Ong, E. S., Segui, P. & Evans, R. M. *Nature* **330**, 624-629 (1987).
3. Petkovich, M. *et al. Nature* **330**, 444-450 (1988).
4. Nusslein-Volhard, C. & Wieschaus, E. *Nature* **287**, 795-801 (1980).
5. Nauber, U. *et al. Nature*, this issue.
6. Giguere, V., Hollenberg, S. M. & Evans, R. M. *Cell* **46**, 645-652 (1986).
7. Akam, M. *Development* **101**, 1-22 (1987).
8. Rushlow, C. *et al. Genes Dev.* **1**, 1268-1279 (1987).
9. Ausubel, F. *et al.* (eds) *Current Protocols in Molecular Biology* (Greene Publishing Associates and Wiley Interscience, New York, 1987).
10. Arriza, J. A. *et al. Science* **237**, 268-275 (1987).
11. Cavener, D. R. *Nucleic Acids Res.* **15**, 1353-1361 (1987).
12. Devereux, J., Haeblerli, P. & Smithies, O. *Nucleic Acids Res.* **12**, 387-395 (1984).
13. Weinberger, C. *et al. Nature* **324**, 641-646 (1986).
14. Hollenberg, S. M. *et al. Nature* **318**, 635-641 (1985).
15. Tautz, D. *et al. Nature* **327**, 383-389 (1987).
16. Campos-Ortega, J. & Hartenstein, V. *The Embryonic Development of Drosophila melanogaster* 9-84 (Springer-New York, 1986).
17. Ilafen, E. & Levine, M. in *Drosophila, A Practical Approach* (ed. Roberts, D. B.) 139-158 (IRL, Oxford, 1986).
18. Jorgensen, E. M. & Garber, R. L. *Promega Notes* **10**, 2-5 (1987).

An eIF-4A-like protein is a suppressor of an *Escherichia coli* mutant defective in 50S ribosomal subunit assembly

Kayoko Nishi*†, Francoise Morel-Deville‡, John W. B. Hershey‡, Terrance Leighton§ & Joachim Schnier*†§||

* Max-Planck-Institut für Molekulare Genetik, Abteilung Wittmann, Ihnestr. 73, D-1000 Berlin 33, FRG

‡ Department of Biological Chemistry, School of Medicine, University of California, Davis, California 95616, USA

§ Department of Microbiology and Immunology, University of California, Berkeley, California 94720, USA

The assembly of ribosomes in bacterial cells is a complex process that remains poorly characterized. The *in vitro* assembly of active ribosomal subunits from purified RNA and protein components indicates that all of the information for proper assembly resides in the primary sequences of these macromolecules. On the other hand, the *in vitro* requirement of unphysiological heating steps suggests that this pathway may not accurately reflect the *in vivo* pathway, and that other proteins may be required. One approach to identify any additional proteins is to isolate second-site revertants of mutants defective in ribosome assembly¹. Ribosomal protein L24 is essential in the assembly of 50S subunits²⁻⁵. We have identified an *Escherichia coli* gene, *srnB*, that, when expressed at high copy number, can suppress the effect of a temperature-sensitive lethal mutation in L24. The *SrmB* amino-acid sequence

Table 1 RNA-dependent ATP hydrolysis by *SrmB* protein

Activator	<i>SrmB</i>	eIF-4A
None	0.2	2.1
Poly(A)	374	
Poly(U)	124	8.6
tRNA	101	
rRNA	342	
R17 RNA	447	
M13 dsDNA	29.5	
M13 ssDNA	67.1	

The incubation conditions for the RNA-dependent ATPase assay are as previously described^{11,13} except that 5% ammonium molybdate in 2M sulphuric acid was used for the extraction. The reactions were carried out in the presence of 100 µM [γ -³²P]ATP (5,000 Ci mmol⁻¹), 1.7 µg *SrmB* protein or 2.4 µg eIF-4A, and 0.1 A₂₆₀ unit of RNA or DNA. The activity was measured by the release of [³²P]phosphate reported as fmol of ATP hydrolysed per s per µg protein. A background of [³²P]phosphate released in the absence of protein and RNA was 3.3 fmol s⁻¹ and was subtracted from each value. ds, double-stranded; ss, single-stranded.

has sequence identity with mouse translation initiation factor eIF-4A and with the human nuclear protein, p68. The purified *SrmB* protein is a nucleic acid-dependent ATPase, like eIF-4A, but can also bind RNA in the absence of ATP and other auxiliary protein factors. The RNA dependent ATPase activity of *SrmB* suggests that like, eIF-4A, it could be involved in specific alterations of RNA secondary structure.

Analysis of pseudo-revertants derived from the ribosomal protein L24 temperature-sensitive (ts) lethal mutation established that the mutant protein is unstable and that a defective interaction of the protein with 23S rRNA was the basis for lethality^{6,7}. We describe here further intergenic suppression experiments using the ts ribosomal assembly defect caused by a missense mutation in the *rplX* gene for ribosomal protein L24. Random cloning and overexpression of *E. coli* chromosomal DNA fragments in combination with a selection for temperature-resistant growth of the *rplX*19 mutant was chosen to isolate possible interacting or compensating genes. Chromosomal DNA from *E. coli* was partially digested with *Sau*3A and ligated into plasmid pACYC184 (ref. 8). The vector has a copy number of about 10 to 20 plasmids per cell, which results in overexpression of any gene containing contiguous transcription start sites. The L24 mutant KNS19 harbouring the *rplX*19 mutation was transformed and chloramphenicol-resistant colonies capable of growth at 42 °C were selected. After two days incubation, two clones were obtained whose plasmids contained DNA fragments with overlapping inserts. Maxicell and complementation analyses using 5' and 3' end deletion mutants led to the identification of a gene (*srnB*) which codes for a protein of relative molecular mass (*M_r*) 50,000 (50K) (*SrmB*). The DNA sequence of *srnB* was determined and an open reading frame encoding a protein of *M_r* 50,158 was identified (Fig. 1a).

The primary structure of the protein was compared with sequences in the DNA databank 'Genbank' (program 'tfastn'). Significant sequence identities were found only with eukaryotic translation initiation factor eIF-4A (ref. 9), which was recently described to be a homologue of human nuclear protein p68 (ref. 10). Comparisons of the protein sequences of *SrmB*, eIF-4A and p68 are shown in Fig. 1b. The *SrmB* protein exhibits approximately 27% sequence identity with both eIF-4A and the p68 protein. The homology occurs throughout the entire protein sequence, suggesting that all three proteins contain sequences that have been conserved in both prokaryotic and eukaryotic lineages. It is, however, noteworthy that in addition to those residues that are common to all three proteins, the *SrmB* protein shares some homologous regions with only the p68 protein and other regions with only eIF-4A. Similarly, the p68 protein shares

† Present address: Department of Biological Chemistry, School of Medicine, University of California, Davis California 95616, USA

|| To whom correspondence should be addressed.

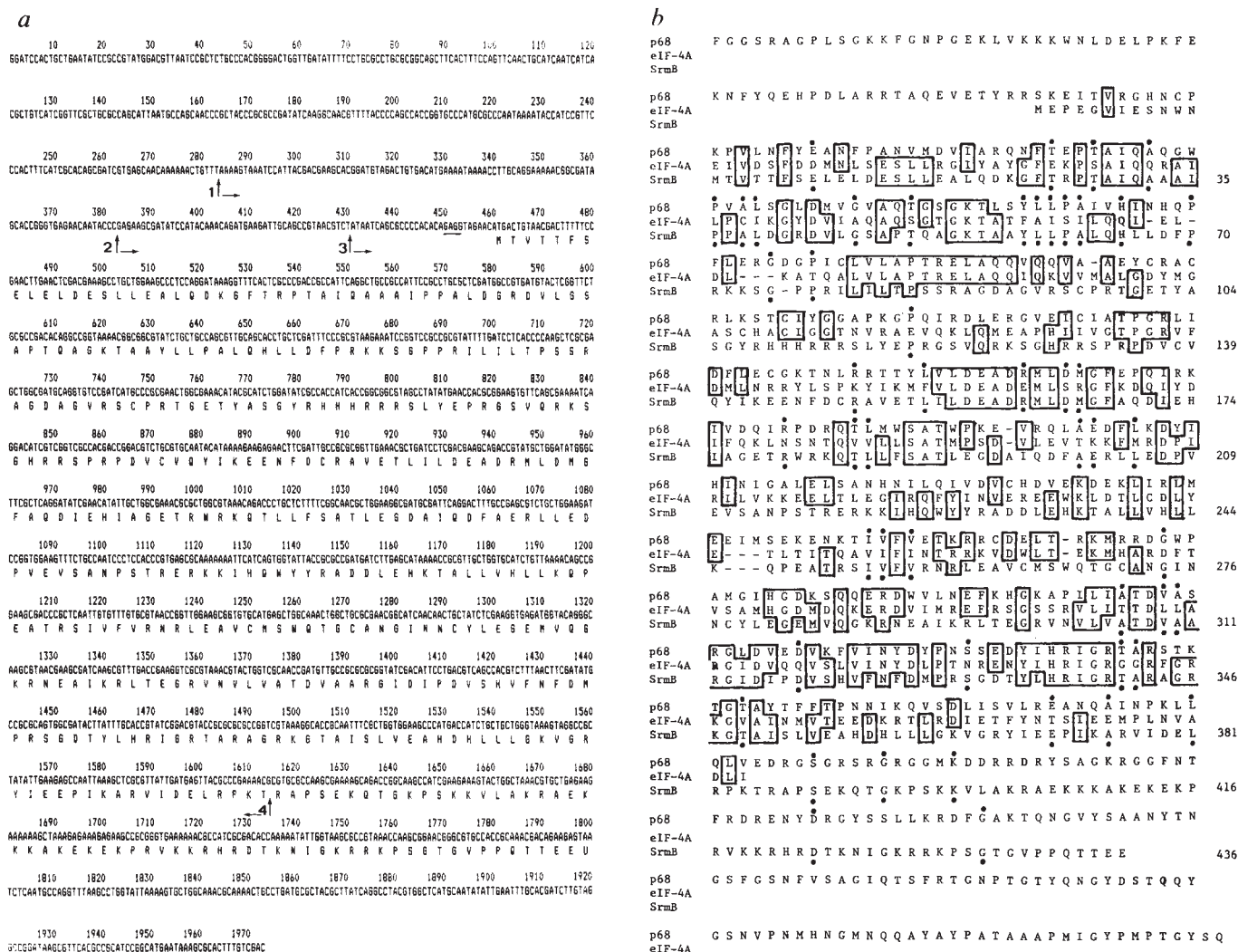


Fig. 1 **a**, DNA sequence and deduced protein sequence of the *srmB* gene. The DNA was sequenced¹⁴ using phage M13 derivatives¹⁵. Both strands were sequenced by shortening a 1.9-kilobase *Bam*HI/*Sal*I DNA fragment with exonuclease *Bal*31. The arrows indicate fragments lacking parts of the 5' end (1, 2 and 3) or of the 3' end (4) that were used for complementation analysis. Only fragments starting at points 1 or 2 could still complement the ts L24 mutant when they were present in plasmid pACYC184 indicating a possible promoter between points 2 and 3. The fragment starting at position 3 was cloned into plasmid pT7-5 for the overproduction of the SrmB protein. A shine-Dalgarno sequence preceding the ATG is underlined. **b**, Amino-acid sequence comparison of *E. coli* SrmB, human p68 and mouse eIF-4A. The alignment of the three proteins used program 'fastn' from Intelligenetics. Identical amino acids between SrmB and eIF-4A, eIF-4A and p68 or all three proteins are boxed; those between SrmB and p68 are marked by dots.

some homologous portions with only eIF-4A. This indicates that all three proteins are derivatives of a common ancestral protein. However, the p68 and SrmB are larger in size than eIF-4A. The alignment of the C-terminal portions of the former two proteins reveals homology in this region, made especially apparent by the frequent occurrence of basic amino acids (Fig. 1b). These data suggest that SrmB might possess functions similar to the other two proteins. Specifically, it seemed important to determine whether or not SrmB has nucleic acid-dependent ATPase activity as has been demonstrated for eIF-4A (ref. 11).

To compare the properties of the SrmB protein with eIF-4A, the protein was purified. A gene fusion of the *srmB* gene was constructed in the T7 promoter expression vector pT7-5 to increase the cellular content of SrmB (Fig. 1a, fragment 3). The overproduction of the SrmB protein caused a reduction in growth rate, but not a cessation of growth. When subcellular fractions were analysed for the SrmB protein, a considerable amount was found associated with the insoluble membrane fraction. For the described purification (see Fig. 2 legend for

details), only the soluble protein present in the S100 fraction was used, which was estimated to contain SrmB as 5–10% of the total cellular protein. About 0.9 mg of essentially homogeneous SrmB was obtained from 25 g (wet weight) of cells. During the protein purification a major difference in the behaviour of the SrmB protein and eIF-4A was noted. Although eIF-4A does not bind to phosphocellulose resin, the SrmB protein bound so strongly that it was one of the last proteins to be eluted from the resin by increasing concentrations of KCl. This unusual binding property could be due to the prevalence of basic amino acids at the C-terminal end of SrmB which is absent in the case of eIF-4A.

The biochemical function of SrmB was tested by measuring its ATPase activity and binding to poly(A). No ATPase activity was detected in the absence of nucleic acid (Table 1). However, in the presence of poly(U) there was considerable activity. The ATPase activity varied in direct proportion to the concentration of poly(U), a property similar to that of eIF-4A (ref. 11; data not shown). Other RNAs such as poly(A), bacteriophage R17 and *E. coli* ribosomal RNA stimulated the ATPase activity to

Fig. 2 SDS-PAGE of the purified SrmB protein. The SrmB protein was purified as follows. *E. coli* strain JS119 carrying *srmb* behind the T7 promoter was grown and SrmB overexpressed. Cells (25 g) were disrupted by Alumina grinding in TM-buffer (10 mM Tris-HCl, pH 7.4; 10 mM MgCl₂; 7 mM 2-mercaptoethanol; 2 mM phenylmethane-sulphonyl fluoride) and the clarified supernatant (S30) was brought to 1 M NH₄Cl before pelleting ribosomes. The S100 supernatant fraction was diluted with four volumes water, then adsorbed to phosphocellulose in P buffer (20 mM K-phosphate, pH 7.5; 7 mM 2-mercaptoethanol; 0.5 mM phenylmethane-sulphonyl fluoride). Protein was eluted with KCl washes of 0.35 M, 0.5 M and 1.0 M in P buffer, and the latter two fractions containing SrmB were concentrated by precipitation with ammonium sulphate. The protein precipitates were dissolved and dialysed against H buffer (20 mM HEPES, pH 7.5; 7 mM 2-mercaptoethanol; 0.5 M phenylmethane-sulphonyl fluoride) containing 50 mM KCl and fractionated by chromatography on a 1-ml MonoS column with a Pharmacia fast protein liquid chromatography system. Following elution with a 50–500 mM KCl gradient in H buffer, SrmB protein was identified by SDS-PAGE¹⁶. The amino-acid composition of the purified protein agreed with the protein sequence deduced from the DNA sequence. Analysis of an aliquot of the pooled fractions is shown. Relative molecular mass markers are shown on the right.

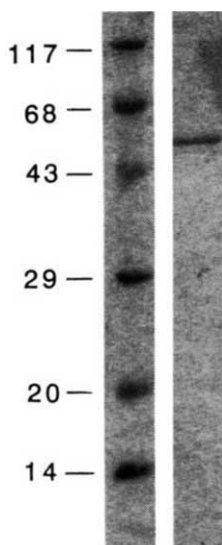


Table 2 Binding of the SrmB protein to poly(A)

[SrmB] (μg)	ATP/Mg ²⁺	Bound [³ H] c.p.m.	poly(A) pmol
1.7	–	7,297	0.93
1.7	+	3,769	0.48
3.5	–	10,913	1.38
3.5	+	8,148	1.03

SrmB binding to poly(A) was determined by the retention of [³H]poly(A) (574 Ci mol⁻¹ of nucleoside residue; Amersham) on a nitrocellulose filter as previously described¹¹. The reaction mixtures containing 23,000 c.p.m. of [³H]poly(A) and the indicated amounts of SrmB protein were incubated for 3 min at 37°C in the presence (+) and absence (–) of 2 mM ATP/Mg²⁺. A background value of 365 c.p.m. obtained in the absence of protein was subtracted from each value.

higher levels than poly(U). Interestingly, DNA, especially single-stranded DNA, also activates the ATPase activity to some extent although we cannot exclude minor RNA contaminations. The RNA-dependent ATPase activity of SrmB is more than an order of magnitude greater than that of eIF-4A.

The binding of SrmB to radiolabelled poly(A) was tested by filtration of complexes through nitrocellulose (Table 2). Significant binding was observed in the absence of ATP and this was decreased by the addition of ATP. Although the RNA-dependent ATPase activity is common to both the SrmB protein and eIF-4A, only the SrmB protein can bind RNA in the absence of other proteins, a property that might be due to the basic C terminus. eIF-4B, which facilitates the binding of eIF-4A to RNA¹¹, had no effect on the binding properties of the SrmB protein (data not shown).

eIF-4A was identified as a factor involved in the initiation of protein synthesis, possibly as a protein facilitating RNA unwinding¹¹. The function of the p68 protein in the nucleus is not yet established, but because of its reaction with a monoclonal antibody to the SV40 large T antigen and because of its location in the cell¹⁰, this protein seems to have a different function from eIF-4A. The SrmB protein is identified here as a suppressor of a bacterial ribosomal assembly mutant. Despite differences in functions and some biochemical properties, the SrmB protein and eIF-4A have at least one important characteristic in common: the RNA-dependent ATPase activity. The high homology among the eIF-4A-type proteins suggests additional common properties that are not shared with other RNA-dependent ATP cleaving enzymes such as ρ factor. It could be that these enzymes have similar RNA target sites, share other characteristics relating to the ATPase activity, or catalyse a specific common effect on RNA secondary structure.

What might be the function of the SrmB protein and what is its connection to protein L24? Previously, we obtained suppressor mutations of the *rplX19* mutant in the 23S rRNA in or close to the binding region of protein L24 and concluded that

protein L24 probably destabilizes a favoured structure of naked 23S rRNA⁷. From the results reported here, it seems likely that the SrmB protein is able to bind to rRNA. The SrmB protein may bind to a similar site on the 23S rRNA as the target site for protein L24. Overproduction of SrmB in stoichiometric amounts to 23S rRNA in the L24 mutant may stabilize the L24 structure in association with assembly at the non-permissive temperature. Alternatively elevated levels of SrmB protein could directly substitute for the assembly function normally provided by L24. This, however, is rather unlikely, because the plasmid containing the *srmb* gene did not complement the growth of a mutant lacking protein L24 (K. N. and J. S., unpublished results). Another possibility is, therefore, that the SrmB protein binds to different regions of 23S rRNA than those occupied by L24 during ribosomal assembly and protects an unstable assembly precursor molecule from degradation. There is evidence for such a defective particle in the L24 ts mutant (30S–32S particles)¹². Whether this precursor is already a partially degraded particle is not clear, as the turnover of several ribosomal 50S proteins in the L24 mutant at the non-permissive temperature is quite high^{6,12}. A stoichiometric amount of SrmB to the precursor particles would, therefore, be required, explaining the gene dosage effect. There is also the possibility that mutant protein L24 itself is protected from protease digestion by the overproduced SrmB protein. In any case, the properties of the SrmB protein suggest that RNA, possibly rRNA is the major target site of the protein. At present we have no evidence that the SrmB protein in wild-type cells also participates in ribosome assembly. One possible function of the SrmB protein however, could be to support the conversion of short-life ribosomal precursor particles to mature 50S subunits by protecting certain regions of the RNA or destabilizing secondary structures.

We thank Dr H. G. Wittmann for support, Sabine Thamm and Susan MacMillan for excellent technical assistance and Dr S Tabor for plasmid pT7-5. This work was supported by grants from the Deutsche Forschungsgemeinschaft, the NIH, the NSF and by a Feodor Lynen fellowship.

Received 22 August; accepted 25 October 1988.

- Schnier, J. & Nishi, K. *Meth. Enzym.* **164**, 706–709 (1988).
- Cabezon, T., Herzog, A., Petre, J., Yaguchi, M. & Bollen, A. *J. molec. Biol.* **116**, 361–374 (1977).
- Nowotny, V. & Nierhaus, K. H. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7238–7242 (1982).
- Nishi, K., Dabbs, E. R. & Schnier, J. *J. Bact.* **163**, 890–894 (1985).
- Nishi, K., Mueller, M. & Schnier, J. *J. Bact.* **169**, 4854–4856 (1987).
- Nishi, K. & Schnier, J. *Molec. Gen. Genet.* **212**, 177–181 (1988).
- Nishi, K. & Schnier, J. *EMBO J.* **5**, 1373–1376 (1986).
- Chang, A. C. Y. & Cohen, S. N. *J. Bact.* **134**, 1141–1156 (1978).
- Nielsen, P. J., McMaster, G. K. & Trachsel, H. *Nucleic Acids Res.* **13**, 6867–6880 (1985).
- Ford, M. J., Anton, I. A. & Lane, D. P. *Nature* **332**, 736–738 (1988).
- Abramson, R. D. *et al. J. biol. Chem.* **262**, 3826–3832 (1987).
- Marvaldi, J., Pichon, J. & Marchis-Mouren, G. *Molec. Gen. Genet.* **171**, 317–325 (1979).
- Grifo, J. A., Abramson, R. D., Satler, C. A. & Merrick, W. C. *J. biol. Chem.* **259**, 8648–8654 (1984).
- Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. *J. molec. Biol.* **143**, 161–178 (1980).
- Messing, J., Crea, R. & Seeburg, P. *Nucleic Acids Res.* **9**, 309–321 (1981).
- Laemmli, U. K. *Nature* **227**, 680–685 (1970).

Sequencing proteins from acrylamide gels

T.E. Kennedy, K. Wager-Smith, A. Barzilai, E.R. Kandel and J.D. Sweatt

The analytical power of high-resolution two-dimensional gel electrophoresis has been coupled with molecular cloning techniques to allow the sequencing of proteins directly from preparative gels.

TWO-DIMENSIONAL polyacrylamide gel electrophoresis has proven to be a powerful analytical tool for detecting and separating proteins. The utility of the approach is significantly enhanced when it is combined with a means to obtain the amino acid sequences of proteins detected on the gels. We have recently developed a method which allows the amino acid sequences of proteins excised directly from Coomassie blue-stained preparative 2-dimensional gels to be obtained¹.

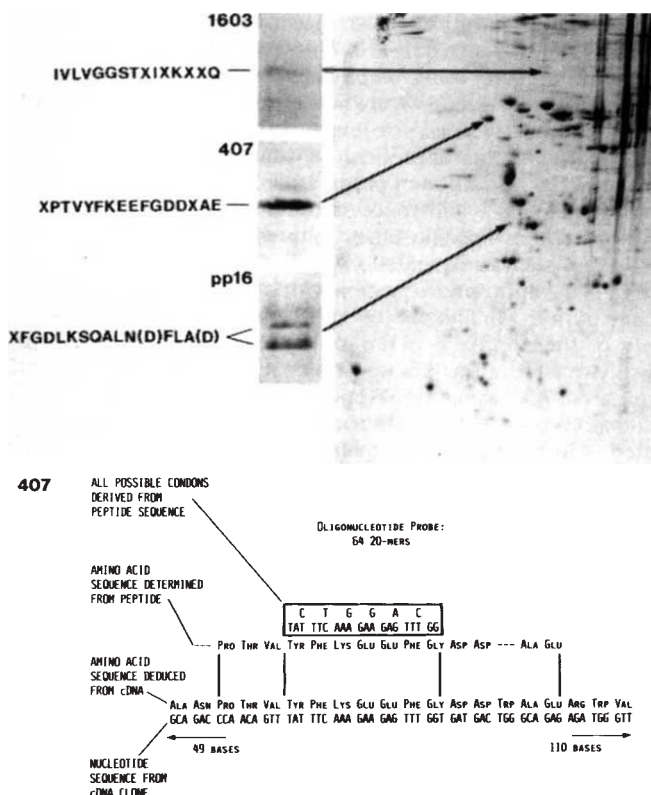
Our method combines established analytical and preparative procedures²⁻⁴. The protein to be sequenced is first isolated by excising it from a Coomassie blue-stained 1- or 2-dimensional gel. The protein is then digested *in situ* with *Staphylococcus aureus* V8 protease, and the resulting peptides are electrophoresed on a second polyacrylamide gel¹. Lastly, the peptides are transferred to a polyvinylidene difluoride membrane (Immobilon, Millipore, Bedford, Massachusetts, as first described by Matsudaira⁵), visualized by Coomassie blue staining, excised, and sequenced using an automated gas-phase sequencer.

The method efficiently combines the following four advantages. First, it is technically straightforward and inexpensive, using reagents and conventional equipment available in most biochemistry laboratories. Second, when combined with automated gas-phase sequencing, it is sensitive and generates amino acid sequence information from starting amounts of protein of as little as 1–10 µg. Third, an internal peptide sequence can usually be obtained because the use of the protease provides a convenient, practical means for sequencing proteins with blocked amino termini. Finally, the use of 2-dimensional gels allows one to purify a particular protein in just a few steps, starting from a crude cell extract and proceeding to an amino acid sequence.

Cellular state studies

Two-dimensional gel analysis is particularly useful for analysing changes in cellular state that result from alterations in many cellular proteins, such as that which occurs following exposure to certain transmitters, second messengers, growth factors, and transforming agents. We developed this method to extend experiments in *Aplysia* in which we found that learning leads to extensive changes in cellular state, as reflected by the pattern of protein phosphorylation and synthesis⁶⁻⁷.

Fig. 1 Top, Coomassie blue-stained preparative 2-dimensional gel of *Aplysia* total central nervous system homogenate, showing amino acid sequence obtained from proteins 1603, 407, and pp16, digested, blotted and sequenced as described in the text. Bottom, strategy for cloning cDNAs corresponding to protein 407. A pool of oligonucleotides containing all possible coding sequences of seven residues of the peptide was constructed. Clones were obtained by screening λgt10 libraries with this pool of oligonucleotides as a probe. The identity of the clone was confirmed by examining the correspondence of codons outside the region probed with the flanking amino acid sequence determined from the peptide.



We illustrate here the utility of the method by focusing on one of these proteins, *Aplysia* protein 407, which can readily be identified on Coomassie blue-stained preparative 2-dimensional gels. Beginning with approximately 1 µg of material pooled from eight Coomassie blue-stained preparative gels, we obtained 14 residues of amino acid sequence from one peptide fragment of *Aplysia* protein 407 (Fig. 1). A search of the Protein Identification Resource of the National Biomedical Research Foundation revealed no homologous proteins. But by using a 2-dimensional gel database provided by Protein Databases, Inc. (Huntington Station, New York), we were able to identify and obtain thirteen residues of amino acid sequence for a homologous protein from rat fibroblasts (rat fibroblast protein 425)⁸, illustrating the potential power of the technique for identifying homologous proteins from different species. The amino acid sequence obtained for *Aplysia* protein 407 was next used to guide the construction of oligonucleotides corresponding to all possible coding sequences. Using this oligonucleotide pool as a probe, we screened λgt10 *Aplysia* cDNA libraries and cloned

cDNAs corresponding to *Aplysia* protein 407 (Fig. 1b).

Amino acid sequence information can also be used to identify previously characterized homologous proteins, through the screening of protein sequence databases. For example, we have also sequenced *Aplysia* protein 1603, another late protein that increases its rate of ³⁵S-methionine incorporation following long-term sensitization⁶ (Fig. 1), and found it to be a member of the heat shock 70 family⁸.

The method may also be used to study proteins of low abundance. We have obtained 19 residues of sequence for *Aplysia* phosphoprotein 16, a protein that increases its level of phosphorylation in *Aplysia* sensory neurons following the application of facilitatory transmitter, as shown in Fig. 1.

We estimate that pooling the material from about 10 preparative 2-dimensional gels should provide sufficient material for the partial sequencing of the 200 most-abundant proteins in most crude cell homogenate preparations. A useful general rule is that sequence may be obtained from proteins visible as Coomassie blue-stained protein spots when material from several gels is combined.

Technical hurdles

There are, of course, limitations to the method. Partial pre-purification may be necessary in certain cases involving proteins of low abundance. In addition, we initially encountered difficulty in successfully electroeluting integral membrane proteins into the second, peptide-separating gel. One strategy which proved successful is the use of reducible cross-linkers in the first preparative gel to improve elution. Also, we have only thus far established the method using *Staphylococcus aureus* V8 protease, and we are now trying to utilize other proteases to expand the applicability of this approach. In some instances, the resolution of preparative 2-dimensional gels will necessitate the development of specific electrophoresis conditions (such as expanded *pI* gradient) to obtain the appropriate protein with sufficient purity. Finally, the overall sensitivity of the approach is limited by the sensitivity of the gas-phase amino acid sequencer. By using an Applied Biosystems (Foster City, California) automated 470A gas-phase sequencer and carefully examining each HPLC record, we have generally obtained reliable sequence information from amounts of peptides in the range of 5 pmol.

Abersold *et al.* have also developed a method for obtaining amino acid sequence information from proteins isolated from 2-dimensional gels⁹. One limitation of their approach is the need for the peptides to be separated by HPLC before sequencing. Although Abersold *et al.* have used their method with success, our method has the advantage of not requiring the HPLC peptide separation step.

Taken to its limits, our method should allow the development of a database of amino acid sequence information that is directly correlated with a 2-dimensional gel database. Given the rapid pace of progress in gene cloning, and the prospects for sequencing entire genomes, a protein database would constitute an important additional context in which to place and use newly acquired protein information. □

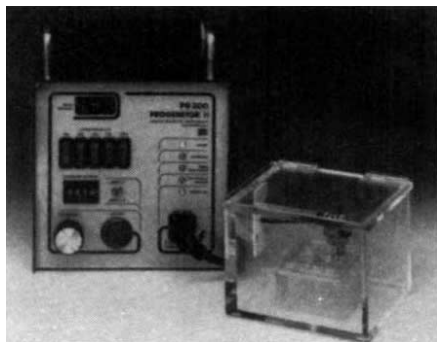
Timothy E. Kennedy, Karen Wager-Smith, Ari Barzilai, Eric R. Kandel and J. David Sweatt are at the Howard Hughes Medical Institute and Center for Neurobiology and Behavior, College of Physicians and Surgeons, Columbia University, 722 West 168th Street, New York, New York 10032, USA.

- Kennedy, T.E. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **85**, 7008-7012 (1988).
- Matsudaira, P. *J. biol. Chem.* **262**, 10035-10038 (1987).
- Cleveland, D.W., Fischer, S.G., Fischner, M.W. & Laemmli, U.K. *J. biol. Chem.* **252**, 1192-1196 (1970).
- Farrell, P.H. *J. biol. Chem.* **250**, 4007-4021 (1975).
- Sweatt, J.D. & Kandel, E.R. (in preparation).
- Castellucci, V.F., Kennedy, T.E., Kandel, E.R. & Goebel, P. *Neuron* **1**, 321-328 (1988).
- Barzilai, A., Kennedy, T.E., Sweatt, J.D., & Kandel, E.R. (in preparation).
- Munro, S. & Pelham, R.B. *Cell* **46**, 291-300 (1986).
- Abersold, R.H., Leavitt, J., Saavedra, R.A., Hood, L.E. & Kent, S.B.H. *Proc. natn. Acad. Sci. U.S.A.* **84**, 6970-6974 (1987).

For gene technology

Products for recombinant DNA studies described this week include a biotin clone detection kit and a system for dissecting yeast asci.

HOEFER Scientific Instruments has just released an updated version of its ProGenetor **electroporation unit** (Reader Service No. 101). The company says the new model PG200 ProGenetor II is designed to deliver low-voltage charges at high capacitance to both plant and animal cells. The unit has a built-in 500 V power supply and five capacitors, which can be engaged individually or in any combination to deliver a maximum discharge of more than 2,770 μ F. Discharges can be programmed for durations of from



From Hoefer in San Francisco, ProGenetor II.

10 μ s to 9,999 ms. ProGenetor II's safety feature does not permit the delivery of a charge unless the electrode is properly in place inside the chamber, preventing accidental operator contact. ProGenetor II comes with an electroporation chamber for holding culture trays and cuvettes, and can be used with both ring and parallel plate electrodes.

Bio-Rad offers a new line of **clone detection kits** based on the method of Kincaid and Nightingale for using biotinylated probes and streptavidin conjugated to alkaline phosphatase to detect positive clones (Reader Service No. 102). Bio-Rad says probes constructed with its Biotin-Blot kits are safer and more stable than ³²P probes, and yield more discrete signals with lower background. The buffers included in the kits are preweighed and pre-mixed, and the enzyme substrates are provided already in solution. The amounts of the kit components have been adjusted so that each is used up at the same rate. The kits are available with two sizes of plaque lift membranes: 137 mm for primary screening and 85 mm for secondary screening. The kits are available with or without specially qualified nylon colony lift membranes.

Singer Instruments has an **ascus dissection system** for yeast genetics studies which automates lineage and tetrad analyses (Reader Service No. 103). The

Singer MSM System consists of a micro-manipulator with ready-made glass needles and an integral microscope with a microprocessor-controlled stage. The system's micromanipulator can be mounted on either side of the microscope stage, for left- or right-hand operation. The console of the system has an LCD display; a joystick controls horizontal movements and a co-axial ring operates the fine vertical drive. Singer says the system has a resolution of 4 μ m and a movement of 115 $\times$ 100 mm, allowing it to accommodate microtitre plates, and both round and square Petri dishes. Alphanumeric coordinates can be keyed in to move automatically to any matrix point. The hinged over-arm of the system allows culture dishes to be interchanged without microtool breakage.

Automated genetics

Eppendorf has a **thermal cycler** for use in enzyme reactions, including the popular polymerase chain reaction (Reader Service No. 104). Eppendorf's MicroCycler can be programmed for temperatures of between 0 °C and 100 °C, and comes with a heat block which accepts 19 1.5-ml microtubes or one which holds 29 0.5-ml tubes. Using the MicroCycler's keypad, up to 99 different time-temperature sub-routines can be defined and stored, linked in any sequence. The MicroCycler is available in both 120 V and 220 V models.

Applied Biosystems has just announced the availability of a new **RNA extraction chemistry** for use with its model 340A nucleic acid extractor, which was introduced last year for the purification of DNA from a wide variety of sample types (Reader Service No. 105). With the new



New chemistry powers the 340A.

chemistry, the model 340A can also purify RNA, using the proteinase K/phenol extraction method. Advantages claimed for automated RNA extraction over

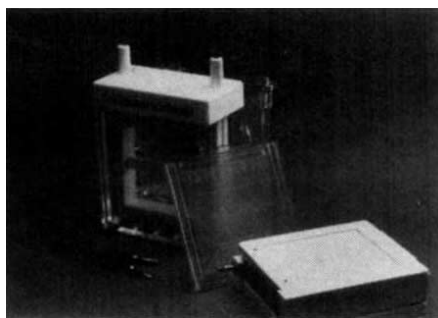
manual methods include high-quality, high-molecular weight RNA consistently, without the chance of sample mix-up or exposure of the user to toxic substances. Samples are loaded directly onto the instrument into eight vessels, where they undergo digestion, organic extraction and alcohol precipitation.

Electrophoresis

Clontech sells a range of DNA and size markers for **pulsed-field gel electrophoresis** (*Reader Service No. 106*). Clontech's Chromo-DNA size markers are available for lambda DNA (50–900 kb), *Candida albicans* (1,000–5,000 kb), *Saccharomyces cerevisiae* (260–1,200 kb), and *Shizosaccharomyces pombe* (1,000–12,000 kb). The markers are provided in 1 per cent low-melting agarose plugs, ready for immediate use. Clontech's Chromo-DNA plugs are intact, chromosomal DNAs suspended in 1 per cent low-melting agarose, ready for digestion with restriction enzymes and subsequent insertion into agarose gels. The DNAs are available for bovine, chicken, human, rhesus monkey, BALB/c mouse and porcine DNA.

Kodak has extended its line of **autoradiography** products to include six new accessory products, including a cassette, an exposure holder, intensifying screens and safelights (*Reader Service No. 107*). The Kodak X-Omatic cassette C-2 is constructed of two vinyl-covered aluminium panels mounted in a polyurethane frame joined by a flexible plastic hinge. The C-2 cassette has a special built-in light lock to eliminate the need for additional light barriers. Kodak's X-ray exposure holder consists of two hinged covers with a metal locking clip to hold the film, specimen and screen in place for autoradiographic exposures. A duplex envelope attached to the front cover folds around the enclosed materials to exclude light. Kodak's phosphor-based screens are mounted on a flexible polyester support with foam backing, and are designed for use with its X-Omatic cassette. Safelight products available from Kodak include a darkroom lamp, an adjustable safelight lamp, and red and brown safelight filters.

The Profile **mini-electrophoresis system** for gel fractionation and electrotransfer is new from Schleicher & Schuell (*Reader Service No. 108*). The compact Profile system uses less than one litre of buffer per run, and can be used to run up to two gels at one time. Both gel fractionation and electroblotting can be performed in the same tank, by simply inserting the electroblotting assembly once the gel run is complete. Schleicher & Schuell says either procedure can be carried out using the Profile system in less than one hour. The Profile system uses Schleicher & Schuell's pre-cast polyacrylamide gels,



See-through electrophoresis: Schleicher & Schuell's Profile system.

which come in a variety of different formats in their own disposable cassettes with a comb. Gels are available for both protein and nucleic acid electrophoresis and blotting procedures; empty gel cassettes are available for pouring custom gels.

Catalogues and services

Pharmacia-LKB's 1989 **electrophoresis catalogue** is now in the mails (*Reader Service No. 109*). The 160-page catalogue contains information on the full line of products from the company's biotechnology division, including reagents, tanks, controllers and blotting systems. The catalogue presents the electrophoresis products developed by both Pharmacia and LKB before their merger together for the first time, and the catalogue contains "everything needed for electrophoresis research except the sample".

Promega offers a **custom oligodeoxy-nucleotide service** for the production of oligonucleotides on a small (0.25 μmol), medium (1 μmol) or large (10 μmol) scale (*Reader Service No. 110*). The custom oligonucleotides are available either purified or unpurified, containing the customer's choice of 5' hydroxyl, titryl or phosphate groups. The oligos are purified by HPLC or gel electrophoresis, and all are claimed to be suitable for direct use in enzymatic reactions. The company will include dI, dU, or dC^{me} residues in the nucleotide sequences for no extra charge.

Oncogene Science's **catalogue and price list** includes such useful products as a kit for the Philadelphia translocation, human oncogene DNA probes, and oligonucleotide probes for oncogenes, growth factors, and *ras* oncogene point mutations (*Reader Service No. 111*). The 13-page catalogue also outlines the company's monoclonal antibodies for DNA tumour virus antigens, growth factor receptors, phosphotyrosine and intermediate filaments, as well as its line of reporter reagents such as biotinylated antibodies, streptavidin conjugates and immunoprecipitation reagents. □

These notes are compiled by Carol Ezzell from information provided by the manufacturers. To obtain further details about these products, use the reader service card bound inside the journal.

ADVERTISEMENTS

ESCHERICHIA COLI TISSUE Culture Incubator

- Safe for AIDS research • Reliable, for In-Vitro fertilization
- Versatile enough for thousands of lab uses

Easy to use – Just flush with gas mixture, seal and place at approp. temp. Perform multiple experiments simultaneously w/out cross contamination. Thousands in use worldwide.

For information contact: **Flow Laboratories, Ltd.** or, **Billups-Rothenberg, Inc.** • P.O. Box 977
Del Mar, CA 92014-0977 USA • (619) 755-3309

Reader Service No.25

O₂/CO₂ micro-Respirometer
Measures gaseous O₂/CO₂ uptake or production with 0.2 $\mu\text{l/h}$ accuracy
Far superior to the Warburg Apparatus

It can be used for measuring O₂/CO₂ production/consumption of insects, fish, small animals, plants, fruits, vegetables, bacterial and tissue cultures, fermentation, oxidation of metals and plastics etc. The results are printed periodically by IBM-PC/XT/AT.

Equipment is also available to measure oxygen consumption in larger animals (mice, rats, pigs, horses etc.).

COLUMBUS INSTRUMENTS
P.O. BOX 44049 COLUMBUS, OHIO 43204 U.S.A.
PHONE: (614) 488-6176 FAX: (614) 276-0529
TLX: 246514

Reader Service No.24

SINGER
INSTRUMENTS

We wish a Happy Christmas and wonderful New Year to all users of SINGER MICROMANIPULATORS

Singer Instrument Co. Ltd.
Roadwater, Watchet, Somerset,
England, TA23 0QL. Tel: Washford (0984) 40226.
Telex: 337492 Corncab G

Reader Service No.50

The ageing of the universities

Richard Pearson

Will the present academic staff of universities around the world be able to cope with the challenges of a new era?

UNIVERSITIES are moving into a new era of static or reducing public funding, a growing relationship with the private sector and consumers (the students), and increased pressure to deliver tangible 'results'.

This all reflects a major change in the philosophy of the decades up to the mid 1970s, when higher education was generally regarded as a Good Thing in an affluent society. During that period, public money flowed into the system to fund the rising cohort of students, resulting from the post-war baby boom and rising educational attainments, thereby completing a virtuous circle as more higher education led to more economic growth.

Young academic faculty flooded into the system for a decade or more with, for example, the number of professors in West Germany quadrupling, faculty numbers in Australia doubling, and in the United Kingdom growing by a quarter. Then in the mid 1970s expansion ground to a halt; hirings came to a virtual standstill, mobility slowed and the population of faculty grew older together. By the mid-1980s the universities had become middle aged. In the United Kingdom the proportion of young staff, aged under 35, fell from 37 per cent in 1972 to 17 per cent in 1985. In Germany it fell from 27 per cent (aged under 40) in 1977 to under 6 per cent in 1983, and in Sweden from over 41 per cent to under 25 per cent. This has led to growing concern about the lack of new talent entering the system, and also the loss of the knowledge of a generation of potential academic researchers who have had to pursue their careers elsewhere.

Fading attraction

The decline in new openings for academics has also lessened the attractions of postgraduate study and research in many countries, leading to an inadequate supply of postgraduate students to meet both current needs and, perhaps more critically, the needs of the next decade. The corresponding growth in the older age ranges (see table) is also seen to have the consequence of lowering staff morale as promotions dry up, creativity stagnates and an apparent wall of retirements builds up which will hit the system at the turn of the century. In West Germany the retirement surge starts rather earlier, with one in five professors expected to retire between 1990 and 1995. These retire-

ments may well coincide with the downturn in supply which was set in train a decade earlier.

The late 1970s did, however, see a growing awareness of these impending problems and a number of initiatives have been established to increase the flow of young researchers into the universities. In the Netherlands an important reform was the restructuring, in 1986, of academic staff into three grades and the establishment of fixed ratios of appointments, with 2.5 senior lecturers for each professor. This complements the introduction of

Percentages of staff aged 50 or over

West Germany	29 (in 1977)	49 (in 1983)
United Kingdom	17 (in 1972)	28 (in 1986)
Norway	30 (in 1977)	38 (in 1985)
Sweden	32 (in 1977)	38 (in 1985)

Source: OECD

AIO or research trainee positions which were introduced in 1984 to provide economic and supervisory support for individuals capable of undertaking advanced research. They involve two- or four-year individually planned programmes which include at least 75 per cent of the time being devoted to research training. Initially 600 posts were created, rising to 5,000 by 1990. In addition two recruitment programmes, targeted at talented PhDs aged 25–40, have been launched providing 350 extra postdoctoral research fellowships. These extra posts are being paid for by a successive reduction in the number of senior university posts and the abolition of temporary appointments, whose numbers are planned to fall from just under 2,000 in 1985 to 485 this year.

In the United Kingdom an early retirement scheme was launched in 1981 which, while it created some extra openings, is believed to have resulted in some staff moving to other well paid positions outside the universities rather than out of the labour force. In addition there has been the 'New Blood' programme of nearly 1,000 extra posts which were successfully targeted at staff aged under 35, with over 90 per cent being under 30, principally in scientific and technological subjects. In addition 110 temporary posts were created under the Royal Society scheme, initially of five years duration but with extensions for up to a further five

years. Finally, additional posts have been created in selected areas such as the new technologies.

Toward the sciences

In West Germany legislation was introduced to facilitate the creation of non-tenured posts in the hope that this would increase staff turnover, while there has also been a redistribution of posts toward the sciences and technologies and away from the humanities and social sciences. Of a more specific nature has been the Heisenberg programme which started in 1977 and is scheduled to run until 1991. Up to 750 five-year awards are available for promising researchers, who were expected to be aged under 33. In the event 543 awards have been made so far across all the disciplines, and although there was a high level of applications the average age of recipients has been 36. Among those completing their awards there has been a strong flow into tenured university or other permanent research appointments. Finally, the Fiebing programme has involved the premature creation of a limited number of tenured university positions between 1985 and 1989 to make extra openings available for young researchers, with a corresponding reduction of posts through retirements in later years.

An interesting early retirement scheme has been introduced in Norway, whereby retirees can not only choose the age at which they leave, but they do not have to leave their home department, maintaining their titles and salaries and being able to devote themselves to full-time research. In addition a number of high priority areas have been targeted in Norway, in the new technologies, to receive a real increase in funding.

The problem of collective ageing is now being addressed in many countries on a selective basis. What is less clear is what is being done to ensure an adequate supply of postgraduates to meet future needs and how academic life can be made more attractive, so that it can match the increasingly competitive private sector as and when recruitment needs soar at the turn of the century. □

Richard Pearson is at the Institute of Manpower Studies, Mantell Building, University of Sussex, Falmer, Brighton, Sussex BN1 9RF, UK.